

UNIVERSITÀ DI PAVIA

FACOLTÀ DI INGEGNERIA

DIPARTIMENTO DI INGEGNERIA INDUSTRIALE E
DELL'INFORMAZIONE

CORSO DI LAUREA MAGISTRALE IN
BIOINGEGNERIA

TESI DI LAUREA

**Predizione dell'aggressività in soggetti autistici:
analisi dell'attività elettrodermica tramite
Random Forest**

Relatore:

Prof. Pietro Savazzi

Correlatore:

Prof.ssa Natascia Brondino

Candidata:

Silvia Barbera

A.A. 2024/2025

Indice

Elenco degli Acronimi	7
Abstract	11
1 Introduzione	13
1.1 Il problema: l'imprevedibilità e l'assenza clinica del dato	13
1.2 I precursori: decodificare il sovraccarico emotivo	14
1.3 Obiettivo del lavoro	14
2 Stato dell'arte	17
2.1 La sfida clinica: imprevedibilità, carico allostatico e la teoria della modulazione dell' <i>arousal</i>	17
2.2 Fisiologia dell'autoregolazione: il sistema nervoso autonomo	19
2.2.1 Il tono di base: iperattivazione cronica e il paradosso dell'ansia	20
2.2.2 L'attività elettrodermica (EDA): non solo stress	22
2.2.3 L'equilibrio cardiaco e il collasso del freno vagale	24
2.2.4 Il ruolo del movimento e l'elaborazione dinamica	26
2.3 Decodificare l'invisibile: indici di allarme e l'equifinalità dei biosegnali	27
2.4 Prevedere le crisi: dalla statistica classica al Machine Learning	29
2.5 Il problema dei falsi allarmi e la creazione dei dati artificiali	30
2.6 Estensione del modello: monitoraggio dello stress in volo	33
3 Dataset	35
3.1 Il contesto clinico e i pazienti coinvolti	35
3.2 La struttura informatica dei file e la codifica dei segnali	35
3.3 La gestione dei dati grezzi: sensori condivisi e sincronizzazione temporale	37
4 L'Acquisizione dei dati	39
4.1 Ambiente e sessioni di acquisizione	39
4.2 I sensori: il braccialetto Empatica E4	40

4.3	Specifiche tecniche del braccialetto Empatica E4	41
4.4	L'infrastruttura software: CareLab ed E4 Connect	42
4.5	L'evoluzione hardware: l'Empatica EmbracePlus e il test pratico	44
4.6	Risultati del test pratico su soggetti sani	45
4.6.1	Caso del Soggetto A: rilassamento e coinvolgimento emotivo	46
4.6.2	Caso del Soggetto B: induzione di stress	47
4.6.3	Considerazioni sul test pratico	49
4.7	Tipologia e significato biologico dei segnali acquisiti	50
4.8	Il protocollo clinico e la definizione delle classi (<i>labeling</i>)	52
5	Pre-processing dei segnali	55
5.1	Dal diario clinico al dato digitale: la definizione del ground truth	55
5.2	Lettura dei file sorgente e sincronizzazione temporale assoluta	56
5.3	Estrazione e processamento delle metriche cardiovascolari (HRV)	59
5.4	Decomposizione dell'attività elettrodermica tramite ottimizzazione convessa	61
5.4.1	Il vincolo di non-negatività e la correzione di offset	66
5.5	Valutazione motoria e continuità del monitoraggio	68
5.6	Analisi multisegnale e risultati del pre-processing	69
5.6.1	Il caso del Soggetto 1	70
5.6.2	Il caso del Soggetto 2	77
5.6.3	Il caso del Soggetto 3	84
6	Estrazione delle feature	91
6.1	Adattamento e calibrazione individuale	91
6.1.1	Definizione della <i>baseline</i>	92
6.1.2	Soglie di validità	94
6.2	Finestratura temporale	95
6.3	Calcolo dei parametri fisiologici	97
6.3.1	L'attività elettrodermica (EDA)	97
6.3.2	Metriche cardiache e motorie	98
6.4	Analisi delle metriche estratte	99

6.4.1	Soggetto 1: attivazione emotiva e capacità autoregolatoria . . .	101
6.4.2	Soggetto 2: il profilo low Responder e l'engagement sensoriale	102
6.4.3	Soggetto 3: Reattività e coinvolgimento attivo	102
7	Modello predittivo Random Forest	105
7.1	Definizione del target clinico	105
7.1.1	Definizione delle soglie di rischio in assenza di crisi	106
7.2	La finestra di previsione e il tempo di intervento	108
7.3	Architettura, addestramento e validazione del modello	109
7.4	Analisi dei risultati	112
7.4.1	Metriche di performance globale	112
7.4.2	Feature importance	113
7.4.3	Il tempo di intervento sui casi reali	116
7.4.4	Il confronto con l'approccio non supervisionato di Change Point Analysis	121
7.4.5	Conclusioni e utilità clinica	122
8	Modello predittivo ibrido e generazione di dati sintetici	125
8.1	L'architettura a due stadi	125
8.2	La generazione dei dati sintetici	126
8.3	Il primo stadio: Change Point Analysis e validazione clinica multimodale	128
8.4	Il secondo stadio: predizione tramite Random Forest	130
8.5	Valutazione delle performance e analisi dei risultati	133
8.5.1	L'identificazione del rischio e il <i>lead time</i>	133
8.5.2	La Matrice di Confusione	135
8.5.3	Capacità discriminativa globale: curva ROC e AUC	136
8.5.4	L'Importanza delle feature	137
8.5.5	Capacità di generalizzazione	138
8.6	Analisi di robustezza al rumore (+6 dB)	139
8.6.1	Ottimizzazione del segnale tramite pre-processing	144
8.7	Analisi di robustezza temporale: variabilità dell'onset	148
8.7.1	Analisi di robustezza temporale del Random Forest	149

8.7.2	Analisi robustezza temporale della Change Point Analysis . . .	150
8.8	Verso l'applicazione reale: il vantaggio del modello ibrido	152
9	Generalizzazione del modello ibrido: dai dati sintetici al contesto reale	155
9.1	Pre-elaborazione avanzata: scomposizione dell'EDA e filtraggio artefatti	156
9.2	Dalle soglie globali alle finestre temporali mobili	157
9.3	Calibrazione intra-soggetto per l'algoritmo Random Forest	158
9.4	Monitoraggio <i>real-time</i> e elaborazione <i>streaming</i> dei dati	159
10	Validazione cross-domain: generalizzazione del modello predittivo sui piloti aerei	163
10.1	Basi fisiologiche e revisione della letteratura	163
10.2	Generazione del dataset fisiologico	164
10.3	Evoluzione dell'architettura ibrida	166
10.3.1	Stadio 1: CPA e baseline personalizzate	166
10.3.2	Stadio 2: Dal Random Forest al RUSBoost	167
10.4	L'elaborazione in tempo reale: masking e filtro causale	168
10.5	Validazione e risultati	170
10.5.1	Analisi dell'importanza predittiva e il ruolo dell'integrazione multimodale	172
10.5.2	Prestazioni e robustezza inter-soggetto	173
10.6	Confronto tra il modello clinico e il modello aeronautico	175
11	Conclusioni e prospettive future	177
11.1	Il significato clinico dei risultati ottenuti	177
11.2	I limiti del lavoro	178
11.3	Sviluppi futuri: verso il transfer learning	178
	Bibliografia	180
	Ringraziamenti	187

Elenco degli Acronimi

ACC	Accelerometria Triassiale (<i>3-axis Accelerometer</i>)
ADC	Convertitore Analogico-Digitale (<i>Analog-to-Digital Converter</i>)
AI	Intelligenza Artificiale (<i>Artificial Intelligence</i>)
ASD	Disturbo dello Spettro Autistico (<i>Autism Spectrum Disorder</i>)
AUC	Area Sotto la Curva (<i>Area Under the Curve</i>)
BLE	Bluetooth a Basso Consumo Energetico (<i>Bluetooth Low Energy</i>)
BOM	<i>Byte Order Mark</i>
BPM	Battiti Per Minuto (<i>Beats Per Minute</i>)
BVP	Volume del Polso Sanguigno (<i>Blood Volume Pulse</i>)
CET	Orario dell'Europa Centrale (<i>Central European Time</i>)
ComEDA	Complessità dell'Attività Elettrodermica (<i>Complexity of Electrodermal Activity</i>)
CPA	Analisi dei Punti di Cambiamento (<i>Change Point Analysis</i>)
CSV	Valori Separati da Virgola (<i>Comma-Separated Values</i>)
cvxEDA	Ottimizzazione Convessa per l'Attività Elettrodermica (<i>Convex Optimization Approach to Electrodermal Activity</i>)
EDA	Attività Elettrodermica (<i>Electrodermal Activity</i>)
FDA	<i>Food and Drug Administration</i>
GABA	Acido Gamma-Amminobutirrico (<i>Gamma-Aminobutyric Acid</i>)
GDPR	Regolamento Generale sulla Protezione dei Dati (<i>General Data Protection Regulation</i>)
HIPAA	<i>Health Insurance Portability and Accountability Act</i>

HR	Frequenza Cardiaca (<i>Heart Rate</i>)
HRV	Variabilità della Frequenza Cardiaca (<i>Heart Rate Variability</i>)
IBI	Intervallo Inter-Battito (<i>Inter-Beat Interval</i>)
IMU	Unità di Misurazione Inerziale (<i>Inertial Measurement Unit</i>)
IoT	Internet delle Cose (<i>Internet of Things</i>)
LB	Limite Inferiore (<i>Lower Bound</i>)
LOO-CV	Validazione Incrociata Leave-One-Out (<i>Leave-One-Out Cross-Validation</i>)
ML	Apprendimento Automatico (<i>Machine Learning</i>)
MV-ASD	Disturbo dello Spettro Autistico Minimamente Verbale (<i>Minimally Verbal - ASD</i>)
NaN	Non un Numero (<i>Not a Number</i>)
NaT	Non un Tempo (<i>Not a Time</i>)
PPG	Fotopletismografia (<i>Photoplethysmography</i>)
PTA	Accelerazione Traslazionale Passiva (<i>Passive Translational Acceleration</i>)
RF	Foresta Casuale (<i>Random Forest</i>)
RMSSD	Radice Quadrata della Media dei Quadrati delle Differenze Successive (<i>Root Mean Square of Successive Differences</i>)
ROC	Caratteristica Operativa del Ricevitore (<i>Receiver Operating Characteristic</i>)
RSA	Aritmia Sinusale Respiratoria (<i>Respiratory Sinus Arrhythmia</i>)
RUSBoost	Sottocampionamento Casuale con Boosting (<i>Random Under-Sampling Boosting</i>)
SaMD	Software come Dispositivo Medico (<i>Software as a Medical Device</i>)

SCL	Livello Tónico di Conduttanza Cutanea (<i>Skin Conductance Level</i>)
SCR	Risposta Fasica di Conduttanza Cutanea (<i>Skin Conductance Response</i>)
SDNN	Deviazione Standard degli Intervalli Normali (<i>Standard Deviation of Normal-to-Normal intervals</i>)
SMOTE	Tecnica di Sovracampionamento della Minoranza Sintetica (<i>Synthetic Minority Over-sampling Technique</i>)
SNR	Rapporto Segnale/Rumore (<i>Signal-to-Noise Ratio</i>)
SNS	Sistema Nervoso Simpatico (<i>Sympathetic Nervous System</i>)
SOP	Procedure Operative Standard (<i>Standard Operating Procedures</i>)
TEMP	Temperatura Cutanea (<i>Skin Temperature</i>)
TNR	Tasso dei Veri Negativi / Specificità (<i>True Negative Rate</i>)
TPR	Tasso dei Veri Positivi / Sensibilità (<i>True Positive Rate</i>)
TSDA	Aumento dei Dati nelle Serie Temporali (<i>Time Series Data Augmentation</i>)
UTC	Tempo Coordinato Universale (<i>Coordinated Universal Time</i>)
VMU	Magnitudo Vettoriale (<i>Vector Magnitude Unit</i>)

Abstract

Vivere con una persona affetta da autismo grave e gestirla, spesso, vuol dire andare incontro all'imprevedibilità del suo comportamento. Infatti, le crisi comportamentali sembrano innescarsi dal nulla, ma in realtà si tratta solo dell'esplosione finale di un sovraccarico emotivo accumulato silenziosamente. A questo si aggiunge l'incapacità di tali soggetti di comunicare verbalmente il proprio disagio. Questo lavoro di tesi nasce proprio con l'obiettivo di riuscire a rilevare ciò che non è visibile dall'esterno, analizzando i segnali biologici e captando da essi gli indici precursori necessari per architettare un sistema di allerta preventivo.

Lo studio ha avuto inizio con l'analisi di dati reali acquisiti da tre soggetti in una stanza multisensoriale tramite i braccialetti Empatica. Fin da subito è emersa una sfida tipica del contesto clinico: l'assenza di crisi reali nei tracciati, dovuta all'efficacia l'intervento preventivo degli operatori o semplicemente perché in quel momento i soggetti non erano predisposti ad alcun crollo emotivo. Tuttavia, è stato possibile approfondire l'analisi e rilevare stati di attivazione emotiva positiva o residui di un attacco di panico pregresso prima che venisse smorzato dal farmaco somministrato; in questo modo è stato possibile individuare la specificità di ogni soggetto e la complessità dei loro comportamenti fisiologici di fronte a determinati stimoli e contesti. Affrontare la mancanza di crisi vere e proprie sotto questa prospettiva ha consentito di studiare i precursori dello stress. Mediante l'uso della piattaforma MATLAB, i segnali sono stati puliti dal rumore e sono stati utilizzati algoritmi avanzati (come *cvxEDA*) per isolare gli impulsi nervosi inviati alle ghiandole sudoripare e al cuore.

Per testare la capacità del sistema di prevedere una situazione di rischio, è stata presa la decisione di creare dei "pazienti virtuali". Sono state costruite delle simulazioni matematiche in grado di riprodurre in modo più fedele possibile la fisiologia di soggetti con ASD sotto stress estremo. Su questi dati è stata addestrata un'architettura ibrida (composta da *Change Point Analysis* e *Random Forest*) capace di rilevare i segnali premonitori prima che la crisi avvenga.

I risultati dimostrano che è possibile ottenere un tempo di anticipo compreso tra 1.5 e 3 minuti prima dell'evento critico, che, in un contesto reale, sarebbe ottimale per consentire un intervento immediato da chi gestisce tali soggetti. Per testare ulterior-

mente la robustezza del modello, è stata estesa la sua applicazione anche al contesto aeronautico, dimostrando che, sebbene lo stress cambi scenario, è possibile adattare il modello secondo opportune ricalibrizioni.

In definitiva, questo studio offre uno spunto metodologico per un'ingegneria predittiva fondata sulla profonda analisi clinica. L'obiettivo è sfruttare i biosensori per intercettare il disagio con un anticipo sempre maggiore, trasformando la tecnologia in un supporto concreto che garantisca a questi soggetti vulnerabili e alle loro famiglie una quotidianità più sicura e serena.

1. Introduzione

1.1 Il problema: l'imprevedibilità e l'assenza clinica del dato

I soggetti con disturbo dello spettro autistico a basso funzionamento, spesso presentano difficoltà nel comunicare il proprio disagio, soprattutto se minimamente verbali (*Minimally Verbal Autism Spectrum Disorder*, MV-ASD). Di conseguenza, questi pazienti accumulano tensione a causa di stimoli ambientali apparentemente innocui, fino a esplodere in episodi di crisi improvvisi che rappresentano un pericolo per il soggetto stesso e per chi lo circonda.

La sola osservazione esterna, purtroppo, non è sufficiente a prevenire tali crisi: i soggetti possono apparire neutri, calmi o sereni mentre il loro sistema nervoso è prossimo al collasso. I biosensori indossabili hanno permesso di superare questa barriera, leggendo i segnali di stress direttamente dal corpo. Tuttavia, il passaggio dalla teoria alla pratica clinica ha evidenziato un ostacolo non trascurabile legato alla rarità degli eventi aggressivi durante le rilevazioni.

Durante l'analisi fisiologica dei dati reali, estratti dai braccialetti *Empatica E4* dei tre pazienti oggetto del presente lavoro, è emerso che i tracciati presentavano informazioni preziose su vari stati emotivi, come l'entusiasmo per stimoli sensoriali o il rilassamento indotto da farmaci o terapie vibroacustiche. L'analisi ha quindi confermato un fenomeno tipico degli studi clinici sul campo: l'assenza di episodi aggressivi acuti nelle finestre di registrazione. In ambito algoritmico, la mancanza fisica di un evento target rappresenta il classico limite degli scenari basati su eventi rari, rendendo di fatto inapplicabile l'addestramento standard di un classificatore supervisionato.

Tuttavia, affrontare questa limitazione intrinseca del dataset non ha compromesso lo scopo della ricerca, ma ne ha elevato il rigore metodologico. Invece di forzare la ricerca di un'anomalia estrema non documentata, l'architettura ingegneristica è stata riprogettata per intercettare e modellare la traiettoria del sovraccarico emotivo, estendendo poi le capacità predittive del sistema tramite simulazioni specifiche fondate su un'attenta analisi clinica eseguita.

1.2 I precursori: decodificare il sovraccarico emotivo

Piuttosto che addestrare l'algoritmo a certificare una crisi già in atto, la metodologia applicata si è focalizzata sull'intercettazione dei suoi precursori. Un episodio aggressivo rappresenta infatti il culmine di una precisa traiettoria di esaurimento delle risorse di autoregolazione. Il sistema realizzato è stato dunque istruito per riconoscere tempestivamente questa serie di alterazioni fisiologiche, identificando uno stato di vulnerabilità acuta definito "sovraccarico emotivo".

Per isolare questo stato transitorio di iperattivazione, il semplice filtraggio dei dati non era sufficiente. Ad esempio, l'attività elettrodermica (EDA), indicatore principale dello stress simpatico, è un segnale complesso che può essere compromesso da fattori come la pelle secca (soggetti *low responders*) o da stati di eccitazione positiva. Per risolvere questa ambiguità, è stata implementata un'avanzata tecnica di ottimizzazione convessa (*cvxEDA*) che ha permesso di scartare il rumore di fondo e isolare i driver neurali puri, ovvero i singoli impulsi nervosi inviati dal cervello alle ghiandole sudoripare.

L'analisi di questi impulsi è stata incrociata con i dati relativi alla variabilità cardiaca (HRV) e all'accelerometria, permettendo di tradurre le teorie cliniche estrapolate in regole matematiche. Conseguentemente, è stato definito il profilo fisiologico del sovraccarico emotivo per ogni singolo soggetto, insegnando al modello a non confondere i movimenti di gioco autostimolatorio (*stimming*) o di interazione sociale con il reale crollo dell'autoregolazione.

1.3 Obiettivo del lavoro

L'obiettivo di questo lavoro è stato strutturato in due macro-fasi complementari. La prima fase ha previsto la massimizzazione del potenziale dei dati reali acquisiti: dopo il preprocessing e la decomposizione dei segnali, è stato addestrato un classificatore Random Forest per prevedere i momenti di massimo sovraccarico fisiologico specifici per ogni paziente.

La seconda fase rappresenta il punto di convergenza metodologica tra due filoni di indagine paralleli e complementari, sviluppati in stretta sinergia. Per superare l'assenza di crisi reali e creare un sistema d'allarme efficiente, si è scelto di ricreare in laboratorio gli scenari mancanti. Dall'integrazione tra regole cliniche e tecniche non supervisionate di *anomaly detection*, è stata sviluppata un'architettura ibrida a cascata. Per testare il sistema, sono stati generati dei "pazienti virtuali" tramite *time series data augmentation*. Le loro reazioni sono state modellate traducendo in parametri matematici le regole cliniche emerse dall'analisi dei pazienti veri e dalla letteratura di riferimento.

L'obiettivo finale del progetto è dimostrare che tale architettura, anche se messa alla prova con forti disturbi e rumore, è in grado di prevedere un'anomalia estrema, garantendo un margine di anticipo (*lead time*) utile nel mondo reale. Nonostante l'enorme variabilità individuale dei soggetti affetti da autismo severo, questo approccio dimostra che è possibile ricalibrare i parametri del modello per adattarlo a diverse casistiche. L'auspicio è che questo studio possa costituire una base metodologica solida per la progettazione futura di dispositivi di allerta *in vivo*, capaci di migliorare concretamente la qualità della vita dei pazienti e dei loro caregiver.

2. Stato dell'arte

2.1 La sfida clinica: imprevedibilità, carico allostatico e la teoria della modulazione dell'*arousal*

All'interno delle strutture di cura e nei contesti familiari, la gestione dei comportamenti aggressivi ed esternalizzanti nei disturbi dello spettro autistico (ASD) rappresenta una delle sfide cliniche più logoranti. L'ostacolo principale non risiede tanto nell'intensità della crisi, quanto nella sua apparente e totale imprevedibilità. Indagando i fattori di rischio su larga scala, Kanne e Mazurek [4] hanno analizzato un campione di 1.380 soggetti, evidenziando un dato controintuitivo: l'insorgenza dell'aggressività non è primariamente correlata al quoziente intellettivo del paziente. I fattori determinanti sono piuttosto la giovane età, i severi deficit nella comunicazione sociale e, in modo critico, una rigidissima resistenza al cambiamento. Queste evidenze epidemiologiche suggeriscono che i *meltdown* (le crisi comportamentali acute) non siano scoppi di rabbia casuali, ma scaturiscano dalla brusca interruzione di pattern ripetitivi e rassicuranti. Si configurano, a tutti gli effetti, come reazioni disadattive a una frustrazione ambientale o a un elevato sovraccarico emotivo che il soggetto non possiede gli strumenti per processare.

Tuttavia, se il quadro clinico generale aiuta i medici a definire il profilo di rischio di un paziente, fallisce quando si tratta di prevenire il singolo episodio acuto in tempo reale. Questo limite diventa invalicabile, specialmente nella popolazione definita "minimamente verbale" (MV-ASD), in cui l'assenza di linguaggio impedisce al soggetto di comunicare il manifestarsi del proprio disagio prima di perdere il controllo. In questi soggetti, l'osservazione esterna del comportamento da parte degli operatori risulta spesso ingannevole. Si tratta del fenomeno che Welch [5] definisce "incongruenza espressiva": un individuo autistico a basso funzionamento può mantenere un'espressione facciale neutra, dondolarsi pacificamente o persino sorridere, mentre all'interno il suo sistema nervoso autonomo sta collassando a causa di un grave sovraccarico

fisiologico.

Per decodificare questo sovraccarico invisibile e superare l'inefficacia della sola osservazione visiva, Goodwin et al. [6] hanno proposto di interpretare il problema spostando il focus dalla psicologia alla fisiologia pura, attraverso la *arousal modulation theory* (teoria della modulazione dell'attivazione). Secondo questo modello clinico, i soggetti a sviluppo neurotipico possiedono un sistema nervoso flessibile, in grado di regolare fluidamente le proprie energie per far fronte agli stimoli. Questo sistema ammortizza gli urti e mantiene l'attivazione fisiologica (*arousal*) all'interno di una sicura finestra di tolleranza. Al contrario, le persone con autismo grave presenterebbero un deficit neurologico intrinseco proprio in questa capacità di filtro. Non riuscendo a smorzare gli stimoli in ingresso (rumori, luci, interazioni), finiscono per vivere in un perenne stato di allarme sensoriale.

A conferma di questa tesi, alcuni di studi basati sull'estrazione dei parametri cardiovascolari hanno scoperto che i pazienti con ASD severo presentano una frequenza cardiaca a riposo significativamente superiore alla norma clinica, accompagnata da una drastica riduzione della varianza. Questa è la prova biologica e tangibile di una cronica iperattivazione simpatica. In termini di ingegneria dei sistemi, è come se il sistema nervoso autonomo di questi pazienti operasse in un costante stato di iperattivazione, al 90% del proprio limite massimo, accumulando quello che viene definito "carico allostatico" [7].

Vivere costantemente al limite delle proprie capacità di sopportazione genera un pericoloso fenomeno noto come *ceiling effect* (effetto soffitto), documentato nei biosignali da Kushki et al. [8]. Dunque, se sottoposti a un nuovo, semplice elemento di stress, come ad esempio un rumore improvviso o un cambio di ambiente, questi soggetti non mostrano la tipica reazione di allarme fasica; i loro tracciati mostrano una risposta fisiologica piatta o "smussata"; ciò vuol dire che il loro corpo non reagisce perché è già saturo. Non avendo letteralmente più "riserve" nervose a disposizione per fronteggiare la nuova richiesta dell'ambiente, il sistema di autoregolazione va in tilt. Lo stress invisibile accumulato si trasforma istantaneamente nella crisi aggressiva visibile agli operatori. L'aggressività non è quindi l'origine del disagio, ma il sintomo estremo di un crollo fisiologico.

2.2 Fisiologia dell'autoregolazione: il sistema nervoso autonomo

Quando si cerca di comprendere le dinamiche che portano a un episodio di aggressività o a un forte sovraccarico emotivo (meltdown) nel disturbo dello spettro autistico, l'istinto più naturale è quello di concentrarsi sul comportamento visibile. Tuttavia, il comportamento esplosivo rappresenta solo l'ultimo anello di una lunga e complessa catena di eventi biologici che sono totalmente invisibili a chi osserva.

Capire dove nasce realmente la crisi prevede una conoscenza del quadro fisiologico e anatomico interno: il sistema nervoso autonomo (ANS). Come suggerisce il nome stesso, questo sistema opera in totale autonomia. Si occupa cioè di regolare tutte quelle funzioni biologiche essenziali che sfuggono al controllo volontario, come ad esempio il battito cardiaco e la sudorazione. In un individuo sano, il mantenimento di questa stabilità interna è affidato a due specifiche componenti del sistema nervoso, le quali agiscono sul corpo provocando reazioni fisiche opposte.

Da un lato, vi è il sistema nervoso simpatico, che si occupa di gestire le situazioni di emergenza e di allerta. Quando l'individuo affronta uno sforzo fisico intenso o percepisce una minaccia, questa componente si attiva immediatamente per mobilitare le risorse energetiche del corpo. In pratica, provoca una risposta fisiologica eccitatoria: innalza la frequenza cardiaca e stimola la produzione di sudore, preparando l'organismo a reagire. Dall'altro lato agisce il sistema nervoso parasimpatico, la cui azione è governata in gran parte dal nervo vago. Questa seconda componente svolge una funzione di recupero: quando la situazione di stress si esaurisce, interviene per disattivare lo stato di allerta. Il suo compito pratico è dunque quello di abbassare il ritmo del cuore e riportare i parametri vitali a una condizione di riposo, favorendo così il rilassamento fisiologico e la predisposizione all'interazione sociale.

In condizioni neurotipiche, questi due sistemi dialogano costantemente per adattare il corpo alle richieste dell'ambiente circostante in modo fluido e naturale. Nei soggetti con autismo severo, come descritto in precedenza, questo bilanciamento fallisce, causando quello stato di "disagio nascosto" (*hidden distress*) [5] per cui un individuo

fisicamente fermo e silenzioso può nascondere una gravissima tempesta autonoma interna.

È proprio a causa di questa profonda disconnessione tra comportamento visibile e fisiologia interna che l'osservazione clinica necessita di un supporto tecnologico. I dispositivi indossabili (*wearable*) diventano lo strumento essenziale per monitorare in modo continuo e oggettivo i due rami del sistema nervoso, fornendo gli algoritmi e i dati necessari per intercettare lo squilibrio autonomo prima che sfoci nella crisi comportamentale.

2.2.1 Il tono di base: iperattivazione cronica e il paradosso dell'ansia

Come introdotto in precedenza, il quadro neurovegetativo di molti soggetti con autismo grave risulta profondamente alterato già in condizioni di assoluto riposo. L'analisi dei dati basali evidenzia infatti una frequenza cardiaca (HR) e un'attivazione elettrodermica (EDA) significativamente più elevate rispetto alla norma clinica. Questo stato di allerta perenne, in cui l'organismo opera costantemente vicino al proprio limite massimo, pone una sfida non solo clinica, ma prettamente ingegneristica.

Kushki et al. [8] hanno analizzato proprio questa dinamica, dimostrando che i pazienti sottoposti a un fortissimo tono simpatico a riposo vanno incontro a un vero e proprio "effetto tetto" (*ceiling effect*). Come illustrato visivamente nella Figura 2.1, l'approccio classico all'analisi dei biosegnali rischia di fallire quando applicato a questi soggetti. In una situazione fisiologica normale (rappresentata dalla curva blu), se una persona affronta un evento stressante imprevisto, il suo corpo reagisce partendo da una linea di base rilassata e producendo un picco fisiologico netto e ben visibile. Nei soggetti con autismo severo, invece, il sistema nervoso sta già operando ai limiti massimi della propria capacità (curva rossa). Di conseguenza, quando si presenta un reale motivo di stress, non c'è più alcun margine biologico per aumentare l'attivazione; il segnale va in saturazione e la risposta non può salire ulteriormente, risultando quindi appiattita o "smussata".

Il rischio è che, agli occhi di un algoritmo addestrato superficialmente, una simi-

le reazioni fisiologiche piatte potrebbero essere facilmente scambiata per un segnale di calma o disinteresse, estremamente lontano dalla verità clinica.

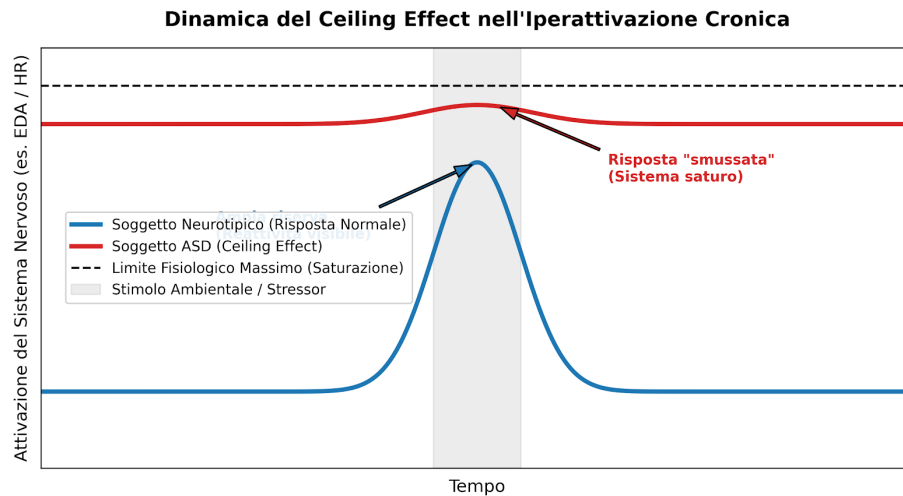


Figura 2.1: Rappresentazione concettuale del *ceiling effect*. A differenza di un sistema nervoso neurotipico, in cui un nuovo stimolo produce un picco di attivazione ben visibile (curva blu), un soggetto in stato di iperattivazione cronica opera costantemente vicino al limite massimo di tolleranza fisiologica. Di fronte a una fonte di stress, l'assenza di riserve biologiche genera una risposta quasi piatta (curva rossa).

Come chiarito dagli studi di Panju et al. [11], questo fenomeno è quasi sempre il sintomo di quella che viene definita "inibizione transmarginale"; nei casi di ansia severa e cronica, il sistema nervoso, non potendo più sostenere quel livello di tensione, crolla e il corpo entra in una sorta di inibizione protettiva, un vero e proprio *shut-down* biologico in cui smette di reagire verso l'esterno per non subire danni peggiori all'interno.

Alla luce di queste dinamiche, appare evidente come un modello predittivo non possa limitarsi a rilevare i semplici picchi di attivazione, ma debba essere istruito a individuare anche i segnali di esaurimento fisico ed emotivo. Il sistema deve essere in grado di comprendere che un tracciato insolitamente piatto, rigido e privo delle normali fluttuazioni non indica affatto che il soggetto sia calmo, ma segnala che l'organismo ha esaurito le proprie risorse di adattamento, trovandosi in uno stato di estrema vulnerabilità che molto spesso precede l'innescò di una crisi comportamentale.

Per completezza scientifica, è doveroso specificare che il disturbo dello spettro autistico, proprio per la sua natura eterogenea, non si manifesta unicamente attraverso l'iperattivazione cronica. Esiste infatti un'altra porzione di soggetti che presenta un

profilo autonomico opposto, caratterizzato da una marcata iporeattività fisiologica. In questi pazienti si osserva dunque un funzionamento opposto: il loro sistema nervoso è cronicamente sotto-attivato; questo fa sì che mostrino una reazione biologica scarsa o nulla persino di fronte a stimoli molto intensi. È una condizione che si traduce spesso nella necessità di ricercare stimolazioni fisiche estreme o dolorose, pur di riuscire a percepire l'ambiente esterno.

La decisione di escludere il profilo iporeattivo da questa analisi, tuttavia, deriva direttamente dall'osservazione dei dati reali. Durante lo studio dei segnali acquisiti da protocollo, non è emerso alcun profilo di vera iporeattività tra i partecipanti. L'unica deviazione ha riguardato un soggetto classificabile come *low responder*, ma i dati suggeriscono chiaramente che non si trattasse di una sotto-attivazione biologica. Piuttosto, il segnale debole era imputabile a limiti fisici di misurazione, come una scarsa aderenza del dispositivo o una secchezza cutanea che ostacolava la lettura del sensore. A causa di queste evidenze pratiche, questo lavoro si concentra maggiormente sulle dinamiche di iperattivazione, per fornire uno strumento capace di intercettare le reazioni aggressive in quei pazienti che vengono travolti da uno stress eccessivo.

2.2.2 L'attività elettrodermica (EDA): non solo stress

L'attività elettrodermica (EDA), che misura le micro-variazioni nella sudorazione della pelle, è riconosciuta come il parametro più sensibile per quantificare l'attivazione del sistema nervoso simpatico. Poiché le ghiandole sudoripare sono controllate esclusivamente da questa specifica componente neurale, ogni fluttuazione rappresenta una scarica nervosa pura e del tutto involontaria. Tuttavia, analizzando la letteratura psicofisiologica più recente e confrontandola con le necessità predittive di un algoritmo, emerge che interpretare questo segnale richiede estrema cautela. L'idea diffusa secondo cui un banale aumento della sudorazione indichi sempre e comunque una crisi imminente risulta essere, dal punto di vista clinico, un'eccessiva e rischiosa semplificazione.

Un contributo essenziale per chiarire questa complessità deriva dallo studio di Greenlee et al. (2024)[10], che impone di scindere l'analisi matematica del segnale EDA in due metriche ben distinte: la frequenza dei picchi e la loro ampiezza. Se il tracciato di un soggetto mostra una frequenza elevata, ovvero produce molti piccoli

picchi ravvicinati nel tempo, non significa che stia andando in una condizione di sovraccarico negativo. Al contrario, questa dinamica fisiologica è risultata fortemente associata a un coinvolgimento sano, attivo e funzionale del soggetto verso l'ambiente che lo circonda (*engagement*), rappresentando a tutti gli effetti un indicatore clinico positivo. Il vero segnale di allerta, quello che un modello predittivo deve imparare a isolare, è dettato invece dall'ampiezza della risposta. In altre parole, quando sul tracciato compare un picco isolato ma di grandissima intensità, ci si trova di fronte a una reazione nervosa anomala, indice predittore di uno stato di profondo disagio.

Oltre all'intensità fisica del segnale, bisogna considerare il contesto in cui si trova il soggetto con ASD, dal momento che lo stesso livello di eccitazione fisiologica potrebbe portare a esiti opposti a seconda di cosa il soggetto sta vivendo in quel preciso istante. Si tratta del cosiddetto "paradosso dell'oggetto" descritto da Prince et al. (2017)[12], secondo cui l'oggetto con cui il soggetto interagisce ne influenza il comportamento. Infatti, i dati clinici dimostrano che se l'EDA subisce un forte innalzamento mentre il soggetto sta giocando liberamente con un oggetto meccanico preferito, non vi è alcun pericolo reale: si tratta di semplice auto-stimolazione eccitatoria (il cosiddetto *stimming*), un meccanismo che viene utilizzato spontaneamente per autoregolarsi. Al contrario, se lo stesso identico aumento del segnale si verifica in una situazione di "desiderio bloccato", come ad esempio quando il soggetto vede un premio ma gli viene imposto di aspettare per ottenerlo, la vulnerabilità cresce a livelli altissimi e il rischio di una reazione aggressiva diventa immediato.

Questa forte dipendenza dal contesto è confermata anche dagli studi di Verneti et al. (2020)[13]. I ricercatori hanno osservato che l'EDA aumenta in modo drastico in risposta a situazioni di chiara frustrazione, fungendo da perfetto precursore fisiologico per l'aggressività reattiva. Eppure, in presenza di uno stimolo percepito dal soggetto come una vera e propria minaccia, gli stessi pazienti possono mostrare una reazione paradossale, caratterizzata da una risposta piatta o da un calo improvviso del segnale (*ipo-arousal*). Questa disattivazione anomala indica uno stato di disinibizione comportamentale che, molto spesso, precede forme di aggressività meno impulsive ma più finalizzate.

Infine, spostando l'attenzione sulla pura prevenzione temporale, lo studio di Fergu-

son et al. (2019)[9] fornisce un *pattern* predittivo fondamentale per istruire un modello di *Machine Learning*. Analizzando le serie temporali, si nota che l'evento aggressivo grave non esplode mai all'improvviso, ma è molto spesso preceduto da quello che viene definito *anticipatory rise*. Si tratta di un innalzamento lento, graduale e costante del livello di base della conduttanza cutanea (la componente tonica del segnale) che inizia a manifestarsi in modo evidente fino a dieci minuti prima dell'esplosione comportamentale.

Oltre a prevedere l'arrivo di una crisi, c'è un altro parametro clinico di fondamentale importanza che il monitoraggio continuo permette di valutare: la reale capacità di recupero del soggetto dopo che l'evento stressante si è concluso. In una situazione ideale, il livello di base del segnale dovrebbe abbassarsi e tornare ai valori di partenza, indicando che l'organismo ha ritrovato il suo naturale equilibrio. Se questo non accade, significa che il corpo non è riuscito a smaltire lo stress accumulato. Questo fenomeno, noto in letteratura come mancata abitudine (*habituation*) e ben documentato anche dagli studi di McCormick et al. (2014)[26], indica che il sistema nervoso del soggetto è ancora completamente saturo. In questo stato di tensione residua, la soglia di vulnerabilità rimane critica e il rischio che si inneschi una seconda crisi aggressiva a brevissima distanza rimane altissimo.

2.2.3 L'equilibrio cardiaco e il collasso del freno vagale

Se l'attività elettrodermica permette di misurare l'intensità della tensione a cui è sottoposto il soggetto, per avere un quadro clinico completo e realmente predittivo, non basta sapere quanto l'ambiente stia stressando il soggetto; è fondamentale capire se il paziente possiede o meno le risorse biologiche interne per gestire quello stress, ovvero l'attività cardiaca.

Un contributo essenziale in merito a ciò è stato fornito da Fenning et al. (2019)[15], nel cui lavoro i ricercatori hanno dimostrato che le crisi comportamentali più gravi (*meltdown*), non sono causate semplicemente da un picco isolato di stress, ma nascono dal fallimento simultaneo dei meccanismi di autoregolazione del corpo. In una situazione di normalità, quando l'organismo va in allerta, il sistema parasimpatico in-

terviene (principalmente attraverso l'azione del nervo vago) per placare il sovraccarico emotivo, rallentare i battiti e riportare la calma.

Fenning ha scoperto che i soggetti autistici vanno incontro alla crisi peggiore quando si verifica una specifica combinazione biologica, definita in letteratura come *pattern* "High-High". In questo scenario, il corpo subisce una spinta eccitatoria massiccia (registrata come un'alta reattività dell'EDA) e, in quello stesso identico istante, subisce un'improvvisa e totale interruzione della modulazione parasimpatica (il cosiddetto *vagal withdrawal* misurabile attraverso un drastico calo della variabilità cardiaca, HRV). In termini pratici, accade che proprio nel momento in cui il livello di stress del soggetto raggiunge il suo apice, l'organismo disattiva improvvisamente l'unico sistema biologico che potrebbe aiutarlo a calmarsi. Quando una tensione così estrema si scontra con la totale assenza di difese interne, la crisi comportamentale smette di essere interpretata come un atto volontario o oppositivo. Si tratta, a tutti gli effetti, del cedimento inevitabile di un sistema nervoso che, non potendo più gestire il sovraccarico, ha letteralmente esaurito ogni sua risorsa.

L'analisi dell'attività cardiaca non serve solo a registrare il picco di stress durante un'emergenza, ma è lo strumento perfetto per cogliere un altro aspetto fondamentale: la fatica silenziosa che precede la crisi. Misurare la frequenza cardiaca media (HR) permette infatti di individuare in modo inequivocabile i segnali dell'ansia anticipatoria. Gli studi di Groden et al. [14] lo spiegano: spesso, nei pazienti autistici, il battito cardiaco si innalza molto prima che l'evento stressante sia cominciato; basta trovarsi in una fase di attesa o ascoltare le istruzioni per un'attività difficile. Agli occhi di un osservatore esterno, il paziente può sembrare calmo, seduto e in silenzio. Il sensore cardiaco, al contrario, rileva che sta compiendo uno sforzo psicologico altissimo solo per processare le informazioni e tollerare quel momento di stallo. Questo innalzamento silenzioso è un dato vitale per il modello predittivo poiché addestra l'algoritmo a comprendere che il soggetto sta consumando le sue difese prima del tempo, e che quando inizierà davvero l'attività critica, sarà già stremato e molto più esposto al rischio di un crollo.

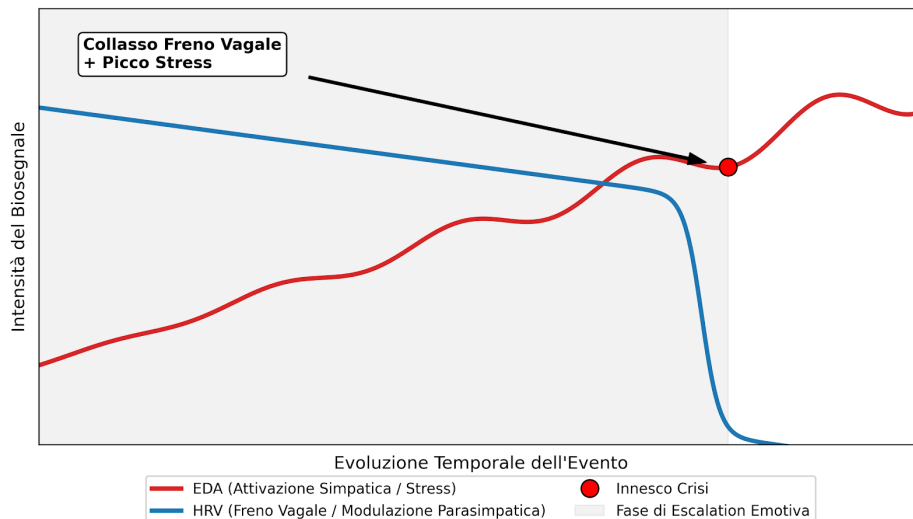


Figura 2.2: Rappresentazione concettuale del pattern “High-High” e del crollo del freno vagale. Il meltdown non è causato semplicemente da un picco isolato di stress. Esso scaturisce dall’incrocio tra una spinta eccitatoria simpatica massiccia (curva rossa dell’EDA) e il contemporaneo collasso totale della modulazione parasimpatica di contrasto (curva blu dell’HRV, *vagal withdrawal*) [15]. Quello evidenziato è il punto di cedimento del sistema nervoso.

2.2.4 Il ruolo del movimento e l’elaborazione dinamica

Per avere una visione completa, l’inquadramento fisiologico composto da EDA e attività cardiaca deve necessariamente incrociarsi con ciò che il soggetto sta facendo fisicamente all’esterno. È qui che entra in gioco il dato meccanico fornito dall’accelerometro. Spesso, chi osserva dall’esterno tende a credere che i classici movimenti ripetitivi dei soggetti autistici (le stereotipie motorie, o *stimming*) siano un modo per calmarsi e abbassare lo stress. Gli studi di Lydon et al. (2013)[27] hanno ribaltato questa convinzione: analizzando il battito cardiaco durante atti di autolesionismo o *stimming*, i ricercatori hanno scoperto che questi movimenti non producono affatto un rilassamento, ma innalzano la frequenza cardiaca o la mantengono costantemente elevata. Molto spesso, infatti, questi soggetti utilizzano il movimento ripetitivo proprio per ricercare attivamente una stimolazione forte, con lo scopo di mantenere il proprio corpo in uno stato di alta attivazione fisiologica. Per un algoritmo che vuole prevedere una crisi, incrociare i dati del cuore e dell’attività elettrodermica con quelli del movimento diventa essenziale per capire il reale contesto di rischio. Come dimostrato da Kanne e Mazurek (2011)[4], i soggetti autistici che mostrano una forte tendenza a comportamenti rigidi e ritualistici presentano il livello massimo di rischio aggressivo

proprio nel momento in cui la loro routine motoria viene bloccata. Se l'algoritmo rileva un movimento ritmico in corso e, improvvisamente, questo *pattern* viene interrotto bruscamente da una richiesta esterna, il sistema deve sapere che la probabilità di una crisi è elevata.

Dall'analisi di tutta questa letteratura clinica si evince che l'aggressività autistica non è assolutamente imprevedibile, come appare agli occhi dei genitori o degli operatori, ma esistono segnali biologici predittivi e misurabili. Tuttavia, l'enorme variabilità di questi segnali rende del tutto inutili i vecchi metodi statistici, che si basano sul calcolo di medie o di linee di base fisse.

Questo perché la fisiologia di questi soggetti oscilla continuamente durante il giorno; dunque, per gestire questa complessità, la ricerca bioingegneristica si sta spostando verso approcci di analisi dinamica e probabilistica, capaci di adattarsi istante per istante al paziente. Strumenti matematici avanzati come i filtri di Kalman (Kushki et al., [16] o la *Change Point Analysis* (Rosenfield et al., [17]) riescono a individuare l'istante esatto in cui la struttura del segnale varia e cambia bruscamente regime. Sono proprio questi metodi a fornire al *Machine Learning* le fondamenta solide per passare dalla semplice osservazione a una predizione tempestiva e davvero affidabile.

2.3 Decodificare l'invisibile: indici di allarme e l'equifinalità dei biosegnali

L'introduzione di biosensori indossabili di grado clinico ha permesso di abbandonare i limiti dell'osservazione visiva per addentrarsi nell'analisi oggettiva dei segnali periferici. Tuttavia, passare dalla teoria medica alla costruzione di un modello predittivo funzionante impone di affrontare sfide computazionali notevoli. Come analizzato in precedenza, l'interpretazione pura e semplice dei biosegnali richiede un'estrema cautela algoritmica. Se si progettasse un sistema basato unicamente sul superamento di soglie di ampiezza fisse, si otterrebbe un modello incoerente, dal momento che la fisiologia autistica, fortemente ancorata al contesto, mostra spesso reazioni che sono l'esatto opposto di ciò che la logica informatica standard suggerirebbe.

La sfida più complessa per un algoritmo di *Machine Learning* applicato all'autismo è dettata da un fenomeno noto in letteratura come "equifinalità". L'aggressività non nasce sempre e solo dal medesimo stato fisico; ad esempio, lo studio di Fenning et al. (2019) [15] ha dimostrato che lo stesso identico comportamento aggressivo grave può esplodere partendo da due condizioni corporee opposte: la prima è legata al profilo *High-High*, che rappresenta il classico *meltdown* da sovraccarico sensoriale ed emotivo. La seconda condizione è il profilo *Low-Low*, che rappresenta lo scenario algoritmico più insidioso, poiché in questa dinamica l'aggressività è più lucida, strumentale e non impulsiva; i dati mostrano una sudorazione quasi assente e un'attività cardiaca rigida e invariabile. Questi non sono affatto indici di tranquillità *Low-Low*, ma i marcatori di un profondo distacco emotivo e non impulsiva; i dati mostrano una totale insensibilità all'ambiente circostante.

Progettare un algoritmo predittivo che si affidi a un solo segnale, ignorando questa complessa dicotomia fisiologica, significherebbe creare un sistema destinato a fallire matematicamente sui soggetti che si intendono proteggere. È per questo motivo che i modelli più recenti richiedono un'architettura multimodale, in cui l'attività elettrodermica, la variabilità cardiaca e l'accelerometria vengano elaborate simultaneamente per fornire il reale contesto del rischio.

Dal punto di vista prettamente metodologico, l'estrazione delle *feature* predittive richiede una rigorosa fase di preelaborazione. Oltre a dover filtrare i disturbi fisiologici, si deve tener conto dell'inaffidabilità dei valori grezzi.

Come sottolineato da Greenlee et al. (2024) [10] nel loro confronto tecnico tra diversi software di elaborazione (NeuroKit2 e Ledalab), i valori assoluti dell'attività elettrodermica cambiano drasticamente a seconda dell'algoritmo utilizzato per la pulizia del segnale. L'algoritmo deve elaborare esclusivamente valori standardizzati (*z-scores*) o differenze relative (*delta*) calcolate rispetto a una baseline dinamica dell'individuo. Solo abbandonando i valori assoluti *ez-scores* abbracciando l'analisi dinamica delle *delta* variazioni, il *Machine Learning* può decodificare l'invisibile e trasformare i segnali biologici in una *prevenzione* tempestiva.

2.4 Prevedere le crisi: dalla statistica classica al Machine Learning

I tracciati analizzati per soggetto con ASD cambiano drasticamente non solo da un individuo all'altro, ma anche nello stesso soggetto a distanza di poche ore. Questa instabilità rende di fatto inutilizzabili gli approcci statistici tradizionali. Se si usassero i classici test T o le ANOVA, basati sul confronto di medie globali e varianze statiche, si andrebbe ad appiattare esattamente ciò che si sta cercando. Infatti, questi strumenti sono nati per confrontare situazioni statiche e, calcolando la media, finiscono per "spalmare" e occultare il singolo, rapidissimo istante di rottura in cui si innesca la crisi. Poiché un *meltdown* non è una condizione fissa ma un'escalation dinamica, cercare di catturarlo con una statistica così rigida farebbe perdere del tutto l'evoluzione temporale del sovraccarico.

Un primo e fondamentale superamento di questo ostacolo si deve all'implementazione di algoritmi basati su *baseline* dinamiche, capaci di adattarsi in tempo reale ai cambiamenti del corpo. L'utilizzo di tecniche di filtraggio predittivo, come il filtro di Kalman impiegato con successo da Kushki et al. [16] sull'intervallo inter-battito cardiaco, ha permesso di isolare le repentine variazioni fasiche legate all'ansia dalle normali e innocue fluttuazioni metaboliche giornaliere (come l'aumento del battito dovuto alla digestione o al cammino). A questo approccio si è presto affiancata l'analisi bayesiana tramite *Change Point Analysis* (CPA), proposta da Rosenfield et al. [17]. Invece di aspettare passivamente che il segnale superi una soglia di allarme prefissata (che cambierebbe da soggetto a soggetto), la CPA calcola istante per istante la probabilità matematica che stia avvenendo un brusco cambio di regime nel tracciato. Questa tecnica si è rivelata estremamente robusta ai picchi isolati (*outlier* fisiologici o artefatti) ed è in grado di intercettare l'instabilità cardio-respiratoria molto prima dell'insorgenza di un attacco di panico acuto, fornendo le basi per l'estrazione di *feature* dinamiche avanzate.

La vera sintesi metodologica per la predizione dell'aggressività e dei comportamenti esternalizzanti nell'ASD è stata formalizzata dal lavoro di Goodwin et al. [2], che ha

segnato il definitivo passaggio al Machine Learning. Il loro approccio si basa sulla segmentazione del flusso continuo dei biosegnali in finestre temporali scorrevoli (*sliding windows*) della durata di 15 secondi. Estrahendo decine di parametri matematici da ciascuna di queste finestre, gli autori hanno dimostrato che l'analisi degli ultimi 3 minuti di tracciato fisiologico contiene informazioni sufficienti per predire con successo un evento avverso nel minuto immediatamente successivo.

Il risultato metodologico più impattante emerso da questo studio non riguarda però i tempi di predizione, bensì la necessità di personalizzazione dell'algoritmo. L'addestramento di modelli di Machine Learning globali (*person-independent*), addestrati cioè su tutti i pazienti contemporaneamente nel tentativo di trovare una regola universale, si è scontrato con un calo delle prestazioni, fermandosi a un'Area Sotto la Curva (AUC) di appena 0.71. Questo fallimento è la dimostrazione matematica dell'equifinalità descritta in precedenza: la struttura del segnale che identifica uno stato di stress acuto per un determinato paziente può coincidere perfettamente con la *baseline* di totale rilassamento di un altro.

Per superare questo limite, la ricerca si è dovuta orientare sui modelli *person-dependent*, in cui un algoritmo specifico viene addestrato esclusivamente sulla variabilità fisiologica di un singolo individuo. Questi modelli su misura hanno quasi raggiunto valori di AUC pari a 0.84, dimostrando una capacità di generalizzazione nettamente superiore. Tale architettura personalizzata è stata recentemente validata su scala multicentrica, dimostrando l'effettiva fattibilità di garantire agli operatori sanitari un tempo di anticipo netto (il cosiddetto *lead time*) fino a 3 minuti prima dell'esplosione aggressiva [3].

2.5 Il problema dei falsi allarmi e la creazione dei dati artificiali

Portare i modelli predittivi fuori dal laboratorio per adattarli alla realtà clinica quotidiana impone di affrontare due ostacoli pratici non trascurabili. Il primo riguarda l'affidabilità: monitorare un paziente in modo continuo espone l'algoritmo a una valanga di falsi positivi, col rischio di generare la cosiddetta *alarm fatigue* e rendere il sistema

inutilizzabile per i terapeuti. Il secondo problema è un paradosso tipico del Machine Learning in ambito medico: per addestrare architetture complesse serve una mole immensa di dati, ma le crisi comportamentali reali sono episodi sporadici, creando un grave vuoto informativo noto in letteratura come *rare event problem*.

Sul fronte operativo, l'ostacolo più insidioso è proprio l'allarme a vuoto. Un sistema di monitoraggio continuo che non sa distinguere l'innescò di un reale *meltdown* da un semplice artefatto meccanico o di movimento finirà per generare allerte ininterrotte. Di fronte a un dispositivo che suona in continuazione senza un reale motivo, la reazione naturale è quella di spegnerlo. Per evitare che la tecnologia risulti invadente, è necessario abbandonare l'approccio di un unico classificatore complesso eseguito ventiquattr'ore su ventiquattro su ogni singolo battito cardiaco, una strategia che comporterebbe solo un insostenibile carico computazionale e un elevato tasso di predizioni errate.

La risposta ingegneristica a questo problema si concretizza nelle architetture a cascata, definite *two-stage anomaly detection*. Il meccanismo, ampiamente supportato dalla letteratura recente [19, 18], si articola su un doppio livello. Un primo algoritmo, leggero e non supervisionato (es. semplici soglie sull'HRV o *change point detection*), funge da filtro in *background*: non esegue diagnosi, ma intercetta solo l'istante in cui il segnale perde stabilità. Solo quando questo filtro isola una finestra temporale sospetta, i dati vengono passati al classificatore vero e proprio (es. *Random Forest*), un modello matematicamente più oneroso il cui compito è confermare l'anomalia o scartarla definitivamente come artefatto. Lavorare con questa struttura ibrida ottimizza drasticamente le risorse della batteria del dispositivo indossabile e abbatte il numero di falsi allarmi. Il secondo limite riguarda la materia prima del Machine Learning: il dataset. Le crisi di aggressività grave sono, per fortuna, eventi poco frequenti. Questa rarità clinica, unita alle rigidissime normative (come il GDPR) che tutelano i dati medici delle categorie vulnerabili, genera il problema delle classi fortemente sbilanciate. L'algoritmo si ritrova ad analizzare migliaia di ore di tracciati fisiologici "a riposo" contro una manciata di minuti di reali "crisi". In uno scenario simile, la rete incappa nel cosiddetto *paradosso dell'accuratezza*: l'algoritmo impara matematicamente che la via più semplice per minimizzare l'errore globale è ignorare del tutto i rari segnali di allarme. Il risul-

tato sarà un modello con un'accuratezza statistica impeccabile su carta, ma totalmente inefficace sul piano clinico, in quanto incapace di intercettare l'evento raro. Persino studi su larga scala condotti in reparti psichiatrici [3] denunciano la rarità degli eventi aggressivi nei dataset come il limite algoritmico principale.

Per superare questo blocco sistemico, la bioingegneria moderna ha dovuto importare le tecniche di *time series data augmentation* (TSDA). Come illustrato da Fawaz et al. [20] e confermato da recenti review [23], la soluzione consiste nel generare parametricamente dati fisiologici sintetici, ricreando *in silico* le risposte nervose di "pazienti virtuali" e bilanciando forzatamente il dataset. Tuttavia, generare dati sintetici non significa banalmente "sporcare" i tracciati aggiungendo rumore statistico casuale. Costruire un segnale biomedico *in silico* impone il rigoroso rispetto dell'anatomia e della fisiologia umana. Ad esempio, nella simulazione della conduttanza cutanea (EDA) non si riproducono curve astratte, ma si impiegano funzioni matematiche precise, come le funzioni bi-esponenziali di Bateman [24], in grado di ricalcare l'esatto processo cinetico con cui il sudore viene secreto dai pori e successivamente evapora dalla pelle. Solo garantendo questo livello di fedeltà biologica [21, 22] è possibile superare i limiti dei dataset tradizionali e garantire un reale incremento delle performance predittive, pur mantenendo la consapevolezza che, in ottica di futura traslazione *in vivo*, si renderà necessaria l'integrazione di tecniche avanzate di deconvoluzione (es. l'algoritmo *cvxEDA* [25]) per pulire ulteriormente i dati grezzi dei pazienti.

Infine, l'adozione di un ambiente di addestramento virtuale travalica la pura necessità matematica legata alle prestazioni algoritmiche, rispondendo a un preciso imperativo etico e normativo. Testare un algoritmo ancora instabile direttamente su pazienti affetti da ASD severo significherebbe bersagliare reparti e famiglie con allerte ingiustificate, esponendoli all'inaccettabile rischio clinico di mancare una crisi acuta. Inoltre, essendo classificati a tutti gli effetti come "software come dispositivo medico" (SaMD), questi algoritmi richiedono severi *stress-test* prima della sperimentazione umana. La simulazione al computer offre lo spazio sicuro necessario per sottoporre il modello a condizioni estreme: simulare la disconnessione improvvisa di un sensore, iniettare rumore di fondo o ricreare scenari limite senza causare alcun disagio ai soggetti. Sotto questa luce, la generazione di biosegnali artificiali smette di essere un mero espediente

matematico per compensare i dati mancanti e diventa un prerequisito metodologico rilevante per la validazione di tecnologie preventive.

2.6 Estensione del modello: monitoraggio dello stress in volo

Il passaggio dalla clinica all'ambito aeronautico richiede una revisione del concetto stesso di risposta allo stress, pur mantenendo intatta l'architettura di base del monitoraggio. Come evidenziato da Rosenblum et al. [43], uno degli errori metodologici più frequenti è trattare la risposta fisiologica come un'entità statica. Infatti, analizzando traiettorie di stress acuto, è emerso che mentre la frequenza cardiaca (HR) risponde quasi istantaneamente, la variabilità della frequenza cardiaca (HRV) presenta una risposta più lenta e con un ritiro vagale significativo solo nella fase centrale dello stress. Questo dato giustifica la necessità di utilizzare modelli dinamici basati su finestre temporali, poiché una media globale del volo nasconderebbe le fluttuazioni critiche del carico cognitivo.

Inoltre, un aspetto importante per la sicurezza del volo è la distinzione tra lo stress cognitivo e quello indotto da fattori ambientali. In aviazione, le accelerazioni traslazionali passive (PTA), come virate, turbolenze o variazioni di quota, sono fattori fisici di stress diretti. Dagli studi di Yu et al. [45] è stato rilevato che l'attività elettrodermica (EDA) reagisce a questi stimoli vestibolari con una latenza di appena 3,3 secondi, aumentando la complessità del segnale (misurata tramite l'algoritmo ComEDA). Queste deduzioni implicano che il sistema di monitoraggio nel *cockpit* deve essere in grado di operare in tempo reale su finestre di pochi secondi per distinguere una reazione di allarme improvviso da un sovraccarico mentale prolungato.

Per garantire l'affidabilità clinica, la ricerca oggi punta sulla *sensor fusion*. In merito a ciò, Coelli et al. [41] hanno dimostrato che l'indice più influente per il rilevamento dello stress cognitivo non è un singolo parametro, ma è dato dalla combinazione multimodale che comprende l'ampiezza fasica dell'EDA, l'entropia del battito e l'attività respiratoria, fondamentale per depurare il segnale HRV dall'influenza del respiro (il fenomeno dell'aritmia sinusale respiratoria, RSA) così da poter isolare la reale attività

del sistema parasimpatico, evitando falsi allarmi dovuti a semplici alterazioni del ritmo respiratorio del pilota durante le manovre.

Infine, l'applicazione pratica a bordo impone l'utilizzo di dispositivi non invasivi ma operativamente robusti. Dalmeida e Masala [42] hanno validato l'uso di *wearable* commerciali per l'estrazione dei parametri HRV, sottolineando una differenza rispetto all'ambito clinico: in contesti di pilotaggio critico, la metrica prioritaria per l'algoritmo deve essere la sensibilità. A differenza dei reparti psichiatrici, dove si cerca di minimizzare i falsi allarmi, in aeronautica il costo di un falso negativo (mancare lo stato di inibizione o lo svenimento del pilota) è fatale, imponendo un diverso bilanciamento matematico del modello predittivo.

Tuttavia, come evidenziato da Wang et al. [44], l'efficacia di questi modelli rimane strettamente legata alla specificità del singolo individuo: ogni pilota processa e manifesta lo stress mentale in base al proprio vissuto e alla propria biologia. Per questo motivo, in perfetta continuità con quanto emerso per i pazienti ASD, i modelli *person-dependent*, addestrati sulla *baseline* fisiologica del singolo operatore, si confermano lo standard ingegneristico assoluto per la futura implementazione dei sistemi di sicurezza *in vivo*.

3. Dataset

3.1 Il contesto clinico e i pazienti coinvolti

Il punto di partenza di tutta l'architettura predittiva sviluppata in questa tesi è costituito dai dati reali. Centro dell'analisi sono stati i tracciati fisiologici di tre soggetti, resi rigorosamente anonimi per ragioni di privacy e identificati d'ora in avanti come Soggetto 1, Soggetto 2 e Soggetto 3, con diagnosi di disturbo dello spettro autistico a basso funzionamento (MV-ASD).

Quando si sceglie di lavorare con dati acquisiti sul campo anziché affidarsi ai classici dataset pubblici da laboratorio, si accetta un compromesso algoritmico preciso: si rinuncia a segnali puliti e perfettamente etichettati per affrontare una fisiologia più complessa e fortemente dipendente dall'ambiente circostante. Pur avendo a disposizione un protocollo di acquisizione esteso, per poter condurre un'analisi solida ci si è concentrati sulle uniche due giornate di acquisizione complete e continue: il 16 e il 18 novembre 2022.

Esaminando a fondo queste ore di registrazione, come anticipato nelle sezioni precedenti, è emerso che non si è mai innescata una vera crisi aggressiva. Inizialmente, l'assenza di questo evento estremo poteva sembrare un ostacolo insormontabile per l'addestramento di un algoritmo. Al contrario, ha finito per definire l'intera direzione del lavoro. Il dataset disponibile è stato sfruttato per indagare ciò che succede prima, tracciando matematicamente l'invisibile accumulo del sovraccarico emotivo, i tentativi di autoregolazione dei soggetti e il modo in cui il loro corpo recupera l'equilibrio.

3.2 La struttura informatica dei file e la codifica dei segnali

Dal punto di vista informatico, i dati grezzi scaricati dai braccialetti si presentavano organizzati in macro-cartelle generate automaticamente dal sistema proprietario di Empatica. Dalla stessa nomenclatura di queste directory (come, ad esempio, la cartella 1668588017_A02724) emergeva la logica di salvataggio del dispositivo: la prima strin-

ga numerica non è altro che l'esatto *Unix Timestamp* dell'istante in cui il braccialetto è stato acceso, seguita dal codice identificativo dell'hardware.

All'interno di queste cartelle si trova il documento di riepilogo `info.txt` generato dal sistema, il quale ha permesso di interfacciarsi con un'architettura di archiviazione molto rigida. Con poche eccezioni, ogni file CSV segue una regola fissa di organizzazione: la primissima riga contiene sempre l'istante di inizio sessione in formato *Unix timestamp* (UTC). Subito sotto, la seconda riga dichiara la frequenza di campionamento espressa in Hertz (Hz). Dalla terza riga in poi, inizia la sequenza temporale dei campionamenti biologici grezzi.

Scendendo nei dettagli dei singoli file a nostra disposizione, vi erano:

- `EDA.csv`: contenente i valori dell'attività elettrodermica misurati direttamente in microsiemens (μS).
- `ACC.csv`: contenente i valori dell'accelerometro triassiale, con i dati degli assi X, Y e Z disposti rispettivamente nelle prime tre colonne. Un dettaglio tecnico fondamentale per la successiva fase di importazione in MATLAB è che il sensore è configurato per un range dinamico di $[-2\text{g}, 2\text{g}]$; di conseguenza, per ottimizzare la memoria del dispositivo, i valori registrati nel file non sono "g" diretti, ma valori scalati in unità di $1/64\text{g}$.
- Il pacchetto cardiaco (`BVP.csv`, `HR.csv`, `IBI.csv`): il file `BVP` riporta l'onda sfigmica grezza della fotoplethysmografia. Da questa, l'algoritmo interno estrae la frequenza cardiaca media (HR) e gli intervalli inter-battito (IBI). A differenza degli altri sensori, il file `IBI.csv` possiede una struttura atipica: non necessita della riga con la frequenza di campionamento. Struttura i dati su due sole colonne, indicando nella prima i secondi trascorsi dall'inizio della sessione e nella seconda la durata esatta di quell'intervallo cardiaco, sempre in secondi.
- `TEMP.csv`: i dati del sensore ottico di temperatura, espressi in gradi Celsius ($^{\circ}\text{C}$).

3.3 La gestione dei dati grezzi: sensori condivisi e sincronizzazione temporale

Di norma, in un esperimento controllato di laboratorio, a ogni soggetto viene assegnata una cartella di salvataggio dati specifica. In questo caso, trattandosi di un protocollo clinico gestito da operatori terzi, la priorità terapeutica era non interrompere le attività della stanza e non infastidire i soggetti con continue manipolazioni dei sensori. Per questo motivo il braccialetto passava spesso dal polso di un soggetto a quello di un altro, senza che la registrazione venisse mai messa in pausa.

Il risultato pratico di questo approccio è stato trovarsi di fronte a file CSV composti da lunghe misurazioni ininterrotte. Inizialmente, dunque, si avevano a disposizione tracciati fisiologici perfetti, ma nessuna etichetta che indicasse chi stesse indossando il dispositivo in un dato momento. Analizzare i file così com'erano avrebbe significato addestrare il modello su un profilo fisiologico inesistente, mescolando i parametri di tre persone diverse, contaminato dai picchi di rumore meccanico generati ogni volta che il sensore veniva sfilato e riallacciato.

L'unica soluzione per frammentare queste registrazioni continue e assegnare la giusta porzione di tracciato al soggetto corretto era incrociare i dati digitali con il diario clinico redatto dagli operatori. Nel farlo, tuttavia, è emerso un secondo problema legato all'allineamento temporale. Come indicato nei file `info.txt`, l'orologio interno dell'Empatica salva i timestamp in formato UTC. Gli operatori, ovviamente, annotavano gli eventi guardando il proprio orologio, basandosi quindi sull'orario locale italiano (che a novembre corrisponde al CET, ovvero UTC+1).

Questa differenza creava uno sfasamento totale sui grafici: se sul diario si leggeva di un tag o di un picco di agitazione alle 17:46, cercando quello stesso orario nel tracciato CSV non si trovava nulla, semplicemente perché per il software quell'evento era stato registrato un'ora prima, alle 16:46. Di fronte a questo scenario, era palese che il dataset grezzo non potesse essere utilizzato direttamente.

Dunque, prima di pensare a qualsiasi analisi predittiva, si è dovuta affrontare una fase di preelaborazione dei dati. È nata la necessità di implementare in MATLAB uno

script su misura che facesse tre cose fondamentali: leggere correttamente i timestamp, compensare l'ora di differenza e segmentare i vettori fisiologici basandosi sugli orari esatti del protocollo clinico fornito. In questo modo è stato possibile estrarre e isolare la vera identità fisiologica di ogni singolo soggetto.

4. L'Acquisizione dei dati

4.1 Ambiente e sessioni di acquisizione

L'analisi dei tracciati fisiologici non può prescindere dal contesto fisico e ambientale di acquisizione; altrimenti significherebbe perdere gran parte del significato clinico del dato. I picchi e le valli osservabili su un grafico dell'attività elettrodermica o cardiaca non sono semplici fluttuazioni matematiche, ma la risposta diretta del sistema nervoso a ciò che accade intorno al paziente. Le sessioni analizzate in questo lavoro non si sono svolte in un laboratorio di ricerca, dove ai soggetti viene solitamente richiesto di restare seduti e immobili; bensì all'interno di un *setting* ecologico e specializzato: una stanza multisensoriale strutturata secondo l'approccio *Snoezelen*.

Si tratta di un ambiente terapeutico progettato su misura per offrire loro un'esperienza immersiva, che gli operatori possono modulare per stimolare o rilassare il sistema nervoso a seconda delle necessità del momento. La stanza metteva a disposizione diversi strumenti per la stimolazione controllata dei canali percettivi: tubi a fibre ottiche e materiali fluorescenti per l'interazione visiva, strumenti per la pressione tattile profonda e, soprattutto, un letto vibroacustico capace di fornire un intenso *feedback* propriocettivo e somatico. Questo approccio terapeutico si fonda sulla creazione di un ambiente privo di richieste prestazionali dirette, un aspetto cruciale per non inficiare i dati fisiologici con l'ansia da prestazione. Le repentine variazioni registrate dai sensori, passando dalla calma indotta dal letto vibroacustico all'eccitazione per un'interazione, riflettono fedelmente quello che la letteratura definisce *social context switch* [5]: l'iperattivazione fisiologica non è dettata da un ambiente fisico ostile, ma dall'impatto della componente sociale e sensoriale.

La profonda conoscenza di queste dinamiche ambientali è stata un prerequisito fondamentale per l'implementazione del codice di pulizia dei dati, permettendo di interpretare l'ingente variabilità riscontrata nei file `.csv`. In un esperimento *standard*, la fase di *baseline* (il riferimento a riposo) è solitamente un segmento piatto e monotono. In questo dataset, al contrario, l'algoritmo si è interfacciato con una fisiologia in continuo e rapido mutamento. Nell'arco della stessa sessione, i sensori hanno re-

gistrato stati di profonda predominanza parasimpatica, indotta ad esempio dalle basse frequenze del letto vibroacustico, seguiti a breve distanza da repentine accelerazioni del battito o picchi di conduttanza, innescati non appena l'attenzione del soggetto veniva catturata dai riflessi luminosi o da un'interazione sociale con gli operatori. Questo contesto dinamico ha imposto un vincolo rigoroso nella fase di pre-elaborazione: non ci si poteva permettere di applicare tecniche di filtraggio troppo aggressive, trattando ogni forte deviazione come se fosse un artefatto di movimento da appiattare. Al contrario, la gestione matematica di questa instabilità è stata calibrata per trattenere le reali risposte biologiche, così da preservare un quadro fisiologico autentico e fedele al modo in cui questi soggetti processano gli stimoli del mondo reale.

4.2 I sensori: il braccialetto Empatica E4

Quando si progetta un sistema di monitoraggio per pazienti con autismo a basso funzionamento, il primo ostacolo di cui bisogna tenere conto è la tollerabilità dell'hardware. Gli strumenti classici della ricerca cardiologica, come le fasce toraciche per l'ECG o gli elettrodi adesivi, risultano di fatto inutilizzabili su questa fascia clinica, poiché significherebbe imporre a questi soggetti un contatto fisico così invadente, con il rischio di innescare un sovraccarico sensoriale tattile e generare una condizione di stress ancor prima di misurarla.

Data l'estrema sensibilità dei soggetti, l'impiego del braccialetto Empatica E4 ha rappresentato la soluzione metodologica ideale. Trattandosi di un dispositivo indossabile al polso, la sua architettura estetica (assimilabile a un comune *smartwatch*) permette di celare la tecnologia e la misurazione medica agli occhi dei pazienti. L'assenza di cablaggi fisici evita che i soggetti si percepiscano "intrappolati", consentendo loro di affrontare le sessioni terapeutiche in modo del tutto spontaneo; un aspetto fondamentale quando si analizzano biosegnali legati alle fluttuazioni emotive.

L'adozione di questa strumentazione ha garantito ai soggetti la massima libertà di movimento all'interno della stanza multisensoriale, permettendo loro di interagire fisicamente con gli operatori senza il rischio di alterare o danneggiare l'hardware. L'aspetto più rilevante, tuttavia, risiede nella totale libertà articolare dei polsi, che ha consen-

tito ai soggetti di ricorrere liberamente allo *stimming* (come il tipico *hand-flapping*). Infatti, per questi soggetti, i movimenti ritmici ripetitivi rappresentano una strategia istintiva di autoregolazione emotiva in caso di sovraccarico. Riuscire a captare la cinematica di questi movimenti di difesa naturale, senza l'interferenza di un equipaggiamento limitante, fornisce all'algoritmo predittivo l'indicazione per cui tali movimenti non devono essere soppressi matematicamente come banale rumore di fondo (*motion artifact*), ma vanno estratti e codificati come un prezioso indicatore di risposta allo stress.

4.3 Specifiche tecniche del braccialetto Empatica E4

Entrando nel dettaglio tecnico, i dati fisiologici analizzati in questo lavoro sono stati acquisiti tramite l'Empatica E4, un dispositivo indossabile di grado medico progettato specificamente per la ricerca continua. L'E4 acquisisce segnali attraverso quattro sensori integrati che lavorano in modo sincrono:

- **Sensore per l'attività elettrodermica (EDA):** Misura le variazioni di conduttanza cutanea attraverso due elettrodi in argento placcato posizionati a contatto con il polso ventrale. Il segnale viene campionato a 4 Hz;
- **Sensore di fotopletismografia (PPG):** Campionato a 64 Hz, utilizza LED verdi e rossi per misurare il volume del sangue periferico, permettendo l'estrazione della frequenza cardiaca media (HR) e dell'intervallo inter-battito (IBI);
- **Accelerometro triassiale (ACC):** Un sensore a 3 assi campionato a 32 Hz, essenziale per rilevare il movimento e filtrare i *motion artifacts*;
- **Termometro a infrarossi:** Misura la temperatura periferica della pelle (campionato a 4 Hz).

Poiché le acquisizioni cliniche all'interno della stanza multisensoriale sono state condotte da operatori terzi, il presente lavoro si è concentrato sull'elaborazione e sulla

classificazione a posteriori di questi dati. Tuttavia, prima di procedere con la stesura dell'architettura in MATLAB, si è ritenuto necessario eseguire delle sessioni di test diretto con la tecnologia Empatica, al fine di validare i flussi di trasmissione.



Figura 4.1: Panoramica dell'hardware, dei sensori integrati e delle specifiche tecniche del bracciale Empatica E4. Immagine estratta dalle specifiche tecniche ufficiali (Fonte:Empatica S.r.l.[46]).

4.4 L'infrastruttura software: CareLab ed E4 Connect

L'estrazione dei segnali dal dispositivo richiede una specifica architettura software di trasmissione e archiviazione. Il sistema di monitoraggio impiegato si articola in tre passaggi fondamentali: l'acquisizione dei segnali dal vivo, il salvataggio automatico su un server online e, infine, il download dei valori numerici sul computer.



Figura 4.2: Dettaglio dell'interfaccia CareLab relativo al monitoraggio della conformità d'uso. La linea temporale evidenzia cromaticamente i periodi di corretta aderenza del sensore, validando a posteriori l'affidabilità meccanica dell'acquisizione.

1. Acquisizione e sincronizzazione locale

Tramite l'applicazione mobile CareLab avviene la prima fase legata all'acquisizione dei dati. Il braccialetto, comunica con il dispositivo mobile tramite protocollo Bluetooth low energy (BLE), garantendo un consumo energetico ottimizzato per sessioni prolungate. Attraverso l'app è possibile visualizzare i tracciati fisiologici in tempo reale, fondamentale per verificare se il sensore sta funzionando bene e, inoltre, consente di inserire dei *marker* temporali per annotare gli eventi salienti.

2. Trasmissione crittografata

In presenza di una connessione internet stabile, avviene l'upload dei pacchetti di dati verso E4 Connect, il portale cloud ufficiale di Empatica. Trattandosi di parametri fisiologici sensibili, la trasmissione avviene in forma crittografata e i server sono strutturati per rispettare rigorosamente le normative internazionali sulla privacy medica (come il GDPR in Europa e l'HIPAA negli Stati Uniti).

3. Estrazione dei dati grezzi

Il terzo e ultimo passaggio riguarda la vera e propria estrazione dei dati. Accedendo alla dashboard web di E4 Connect da un computer, è possibile visualizzare l'intera sessione clinica sotto forma di grafici.

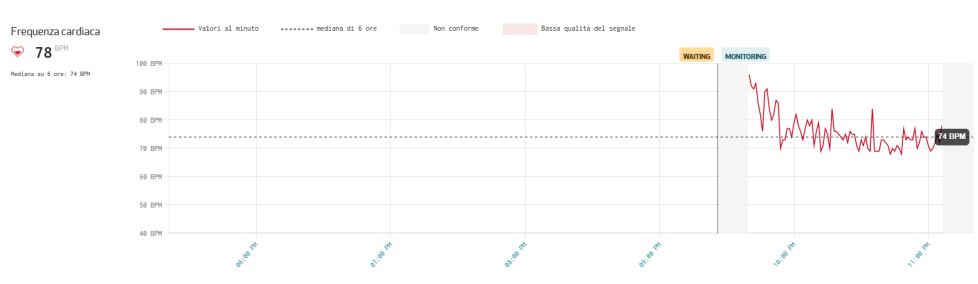


Figura 4.3: Esempio di pre-visualizzazione dei tracciati fisiologici (dettaglio HR) tramite l'interfaccia web del portale CareLab, prima dell'estrazione dei dati grezzi.

Come ampiamente descritto nel capitolo 3, l'architettura intrinseca di questi file generati dal cloud (strutturati su *Unix Timestamp* e frequenze di campionamento differenziate) ha reso inapplicabile una lettura diretta, rendendo strettamente necessaria l'implementazione del codice MATLAB per la corretta associazione cronologica, la

sincronizzazione dei fusi orari (da UTC a CET) e l'allineamento con il protocollo clinico.

4.5 L'evoluzione hardware: l'Empatica EmbracePlus e il test pratico

Per condurre il test pratico, vi è stata l'opportunità di utilizzare l'Empatica EmbracePlus, il dispositivo medicale di nuova generazione (certificato CE classe IIa e approvato FDA 510(k)) che rappresenta l'evoluzione diretta dell'E4, rispetto al quale presenta un'architettura hardware e software rinnovata, con un design più compatto (40 g di peso), uno schermo e-ink "always-on" protetto da vetro Corning® Gorilla® Glass 3 e un'interfaccia a due pulsanti meccanici per il *tagging* degli eventi. Come riportato nella documentazione tecnica ufficiale [50], i sensori integrati hanno subito un importante aggiornamento:

- **Il sensore PPG:** è stato riprogettato utilizzando un array ottico più complesso (4 canali multi-lunghezza d'onda, con l'aggiunta di LED a infrarossi). Questo aggiornamento hardware, unito a nuovi algoritmi di cancellazione adattiva del rumore basati sui dati di movimento, offre una maggiore resistenza agli artefatti, garantendo un'estrazione più pulita e continua dei parametri cardiovascolari anche durante l'attività fisica.
- **Il sensore EDA:** mantiene il posizionamento sul polso ventrale, ma è stato dotato di convertitori analogico-digitali (ADC) ad altissima risoluzione e di un *range* dinamico più ampio. Questo permette di catturare fluttuazioni e risposte galvaniche (SCR) microscopiche del sistema simpatico che i modelli precedenti avrebbero confuso con il rumore di fondo cutaneo.
- **Il comparto motorio:** è passato da un semplice accelerometro a una Unità di Misurazione Inerziale (IMU) a 6 gradi di libertà (accelerometro e giroscopio triassiali), permettendo una mappatura cinematica estremamente più precisa.



Figura 4.4: Rendering ufficiale del braccialetto Empatica EmbracePlus. Si nota il display e-ink integrato e il design ergonomico. (Fonte: Empatica Inc. [47]).

L'innovazione più impattante riguarda tuttavia l'infrastruttura software di esportazione e la gestione dinamica delle frequenze. Mentre il precedente modello E4 acquisiva dati a frequenze rigide (es. 64 Hz per il BVP, 4 Hz per l'EDA), l'EmbracePlus rimodula automaticamente la frequenza di campionamento (fino a 64 Hz) in base al *Sensor Mode* selezionato, ottimizzando così la batteria. Inoltre, la nuova piattaforma *Empatica Health Monitoring* adotta un'architettura di *storage* differente [50]: i dati grezzi ad alta frequenza vengono crittografati in complessi file `.avro`, mentre i file `.csv` resi immediatamente disponibili sulla *dashboard* costituiscono i cosiddetti *Digital Biomarkers*. Questi ultimi sono metriche fisiologiche pre-elaborate dall'algoritmo proprietario e aggregate in finestre temporali fisse di un minuto. Questa differenza architettonica è fondamentale per comprendere la natura a bassa risoluzione dei grafici estratti durante il test.

4.6 Risultati del test pratico su soggetti sani

Per testare la reattività del nuovo dispositivo, l'EmbracePlus è stato applicato su due soggetti neurotipici (Tabella 4.1) sottoposti a stimoli audiovisivi opposti, al fine di ot-

tenere risposte fisiologiche plausibili, adatte al soggetto in esame già interrogato nella fase di definizione del protocollo. Per l'analisi in MATLAB dei *Digital Biomarkers* scaricati, è stato implementato un filtro a media mobile con finestra dinamica pari al 5% dei campioni totali, utile per evidenziare il *trend* fisiologico di base separandolo dagli artefatti transitori.

Tabella 4.1: Riepilogo delle sessioni di acquisizione per il test.

ID Soggetto	Sesso ed Età	Stimolo / Task	Data	Orario (Rome)	Durata
Soggetto A	Uomo, 27	Film autobiografico	12 Aprile 2026	21:30 – 22:53	~83 min
Soggetto B	Donna, 26	Film Horror	10 Aprile 2026	14:30 – 15:27	~57 min

4.6.1 Caso del Soggetto A: rilassamento e coinvolgimento emotivo

Il Soggetto A è stato monitorato per una sessione di 83 minuti durante la visione di un film biografico. L'analisi del tracciato multi-parametrico (Figura 4.5) permette di osservare il progressivo instaurarsi di un profondo coinvolgimento emotivo, caratterizzato da una risposta fisiologica coerente con lo stimolo narrativo.

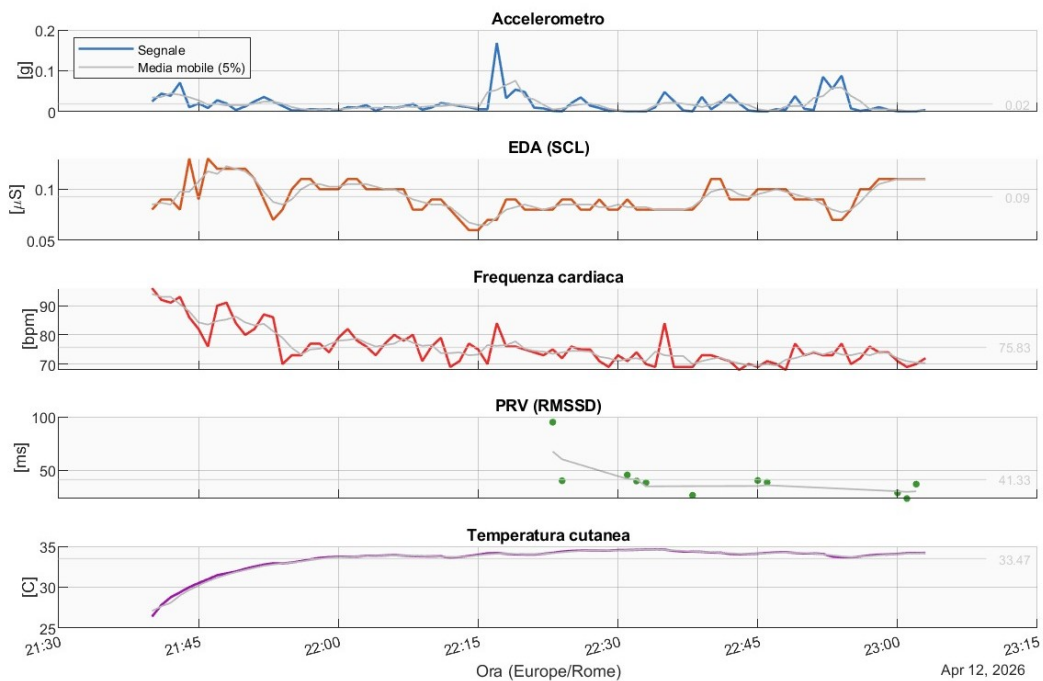


Figura 4.5: Andamento dei *Digital Biomarkers* estratti tramite MATLAB per il Soggetto A durante la visione del film autobiografico.

L'analisi dell'attività motoria (ACC) mostra valori costantemente inferiori alla soglia di 0.05 g per quasi l'intera durata della sessione, confermando lo stato di riposo fisico del soggetto. L'unico scostamento di rilievo è rappresentato da un picco isolato intorno alle 22:18 (circa 0.17 g), riconducibile a un fisiologico aggiustamento posturale del soggetto.

Sotto il profilo cardiovascolare, la frequenza cardiaca (HR) si è mantenuta su una media di 75.8 BPM, valore indicativo di uno stato di calma. Le fluttuazioni del segnale hanno assunto la forma di onde morbide, con picchi rari fino a 96 BPM in corrispondenza dei momenti di maggiore *climax* o commozione narrativa. Al contrario, il grafico a dispersione della variabilità cardiaca (PRV-RMSSD) evidenzia la criticità dei dati aggregati: su 83 minuti di registrazione, risultano validi solo 11 minuti. Questa frammentazione è la conseguenza diretta della sensibilità del sensore ottico PPG, che porta l'algoritmo *cloud* di Empatica a scartare le finestre temporali contrassegnate dai *flag* di errore `worn_during_motion` o `worn_with_low_signal_quality` [50].

L'attività elettrodermica (EDA) ha registrato valori assoluti ridotti rispetto agli standard normativi, con un incremento tonico che è passato da 0.06 μS a un massimo di 0.13 μS intorno alle 21:45. Lo schiacciamento dell'ampiezza fasica è imputabile alla natura aggregata del dato al minuto, che sopprime le variazioni rapide privilegiando il *trend* tonico. Clinicamente, l'ascesa lenta e costante del segnale suggerisce un'attenzione sostenuta e un'empatia crescente verso lo stimolo visivo, escludendo fenomeni di assuefazione passiva.

Il quadro fisiologico è completato dal tracciato della temperatura cutanea, che si è mantenuto stabile su una media di 33.4 °C. L'assenza di crolli termici improvvisi conferma l'assenza di reazioni di allarme o vasocostrizione periferica, convalidando uno stato di elaborazione emotiva profonda ma sicura.

4.6.2 Caso del Soggetto B: induzione di stress

Il Soggetto B è stato sottoposto alla visione di un film horror per una sessione della durata di 57 minuti.

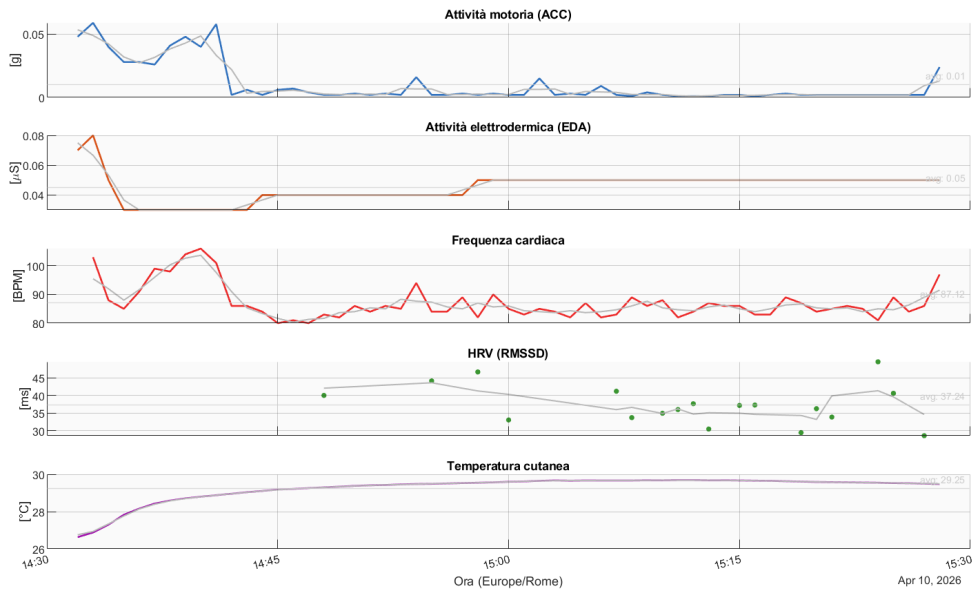


Figura 4.6: Andamento dei *Digital Biomarkers* estratti tramite MATLAB per il Soggetto B durante la visione del film horror.

L'analisi dell'attività motoria (ACC) evidenzia che, dopo una fisiologica fase iniziale di assestamento posturale (fino alle 14:42 circa), il tracciato si stabilizza in prossimità dello zero per il resto dell'esperimento. Questa prolungata immobilità attesta la corretta esecuzione del *task* di visione passiva (il soggetto era sdraiato) e garantisce che le alterazioni rilevate negli altri parametri siano di natura prettamente emotiva e non artefatti da sforzo fisico.

Sotto il profilo cardiovascolare, il tracciato della frequenza cardiaca mostra una *baseline* dinamica piuttosto mossa, con una media di 87.12 BPM. Si evidenziano due momenti di forte attivazione simpatica: un primo innalzamento intorno alle 14:40 e il picco massimo della sessione a 106 BPM subito dopo le 15:00. Quest'ultimo picco è perfettamente sincrono con un micro-movimento registrato dall'accelerometro, evidenziando una chiara reazione traslazionale di allarme del soggetto (il tipico *jump scare*). Anche in questo caso, il calcolo vagale (PRV-RMSSD) risulta severamente frammentato; tuttavia, i rari segmenti validi registrano una netta caduta dai 45 ms iniziali a minimi di 28 ms nelle fasi di tensione, confermando biologicamente il ritiro del "freno" parasimpatico.

Il segnale relativo all'attività elettrodermica (EDA) si è rivelato il più carente e compromesso. Dopo aver raggiunto un picco transitorio di appena 0.08 μ S intorno

alle 14:40, il segnale crolla appiattendosi su valori minimi (intorno a $0.03 \mu\text{S}$) per il resto della sessione, restituendo una media fissa insolitamente bassa ($0.05 \mu\text{S}$). Questa rappresentazione "piatta" è probabilmente dovuta alla pre-elaborazione del *Digital Biomarker*: l'aggregazione a un minuto maschera completamente eventuali risposte fasiche (SCR) di breve durata, restituendo un tracciato tonico smorzato.

Tuttavia, il crollo permanente del segnale richiede un'interpretazione clinica e strumentale più profonda, che apre a due ipotesi non mutuamente esclusive. La prima è fisiologica: il soggetto potrebbe aver sviluppato un rapido effetto di assuefazione allo stimolo visivo; dopo i primi eventi di tensione, le ghiandole sudoripare si sarebbero adattate allo stato di allerta prolungato smettendo di rispondere, sebbene il cuore continuasse ad accelerare. La seconda ipotesi, di natura puramente hardware, suggerisce che il sensore abbia perso la sua efficacia di conduzione a causa della secchezza della cute o di un allentamento del cinturino.

Questa seconda ipotesi (l'errore strumentale) è fortemente supportata dal tracciato della temperatura cutanea, che mostra un improvviso crollo termico passando da $29.5 \text{ }^\circ\text{C}$ a un minimo di $26.6 \text{ }^\circ\text{C}$. Sebbene lo stress acuto induca una forte vasocostrizione periferica (richiamando il sangue verso gli organi centrali e raffreddando il polso), un valore di $26.6 \text{ }^\circ\text{C}$ per la pelle di una persona al chiuso è fisiologicamente inverosimile. È dunque altamente probabile che, in corrispondenza di un movimento di allarme, il braccialetto si sia parzialmente distaccato dalla cute, portando il termometro a misurare la temperatura più fredda dell'aria circostante e azzerando simultaneamente la conducibilità del sensore EDA.

4.6.3 Considerazioni sul test pratico

Affidarsi ai *Digital Biomarkers* aggregati (media al minuto) forniti dai sistemi *cloud* commerciali può risultare insufficiente per l'addestramento di un modello predittivo tempestivo. L'aggregazione temporale appiattisce le dinamiche fasiche ad alta frequenza e l'algoritmo di pulizia proprietario elimina porzioni di segnale (*missing data*) al primo accenno di movimento. Considerando che l'obiettivo di questa tesi è analizzare i segnali di soggetti affetti da ASD, caratterizzati da forte irrequietezza motoria (*stimming*), l'utilizzo di dati pre-aggregati e pre-filtrati comporterebbe la perdita totale delle

informazioni più critiche. Questo giustifica scientificamente la necessità, per il prosieguo del lavoro, di tornare ai dati grezzi (*raw data*) ad alta frequenza acquisiti tramite l'Empatica E4. Solo lavorando sul segnale puro(ad es. 64 Hz e 4 Hz) e applicando un processo specifico di pre-elaborazione in MATLAB, sarà possibile estrarre i precursori che descrivano l'instabilità del segnale senza perderne il valore clinico predittivo.

4.7 Tipologia e significato biologico dei segnali acquisiti

Per comprendere e interpretare correttamente lo stato interno di soggetti incapaci di verbalizzare il proprio disagio, affidarsi a un singolo parametro fisiologico è un errore metodologico noto. Il dispositivo Empatica E4, grazie alla sua natura intrinsecamente multimodale, permette di acquisire informazioni sull'attività del sistema nervoso autonomo nella sua complessa dicotomia (simpatico e parasimpatico).

I dati ottenuti dalle sessioni analizzate permettevano di ricavare serie temporali strutturalmente molto diverse tra loro, salvate in file .csv indipendenti. Di seguito viene presentato nel dettaglio il set di dati a disposizione e la motivazione sia clinica che matematica per cui ogni singolo parametro si è rivelato fondamentale per l'addestramento del modello predittivo:

- **Attività elettrodermica (EDA):** Acquisita tramite due elettrodi posizionati a contatto diretto con il polso ventrale, misura le variazioni microscopiche di conduttanza legate alla sudorazione. Questo è stato in assoluto uno dei segnali più informativi dell'intera analisi. Poiché le ghiandole sudoripare eccrine sono innervate esclusivamente dal sistema nervoso simpatico, l'EDA rappresenta uno specchio puro, affidabile e immediato dell'eccitazione emotiva, dello stress e del cosiddetto *arousal*. Un soggetto autistico in estrema difficoltà può mascherare il proprio stato mantenendo un'espressione facciale neutra o restando in totale silenzio (incongruenza espressiva), ma non possiede alcun controllo volontario o cognitivo sulla propria sudorazione. L'EDA offre dunque l'enorme vantaggio

predittivo di rilevare il sovraccarico autonomico e la frustrazione interna diversi minuti prima che questi si trasformino in una crisi aggressiva esternalizzata.

- **Accelerometria triassiale (ACC):** Il sensore di movimento spaziale lavora misurando le accelerazioni sui tre assi (X, Y, Z). Mentre in gran parte degli studi clinici *standard* l'accelerometro viene impiegato esclusivamente come riferimento per identificare e filtrare i *motion artifacts* dai segnali più delicati, in questo specifico contesto di studio il movimento costituisce un'informazione clinica di rilievo. Infatti, incrociando matematicamente i dati dell'accelerazione spaziale con la frequenza e l'ampiezza dei picchi di conduttanza cutanea, è stato possibile per l'algoritmo comprendere il contesto, permettendo al sistema di discriminare un innocuo e ritmico micromovimento auto-calmante (*stimming*) da una reale agitazione fisica incontrollata che preannuncia l'arrivo di un *meltdown*.
- **Fotopletismografia (BVP) e variabilità cardiaca (HRV):** Attraverso i suoi sensori ottici il dispositivo mappa in continuo l'onda del volume sanguigno periferico (BVP). Da questa lettura ottica, l'algoritmo interno di Empatica estrae l'esatta distanza temporale tra un battito cardiaco e il successivo, salvando questa metrica vitale nel file *IBI.csv (Inter-Beat Interval)*. A partire da questi dati grezzi, il codice ha potuto calcolare parametri non lineari importanti come la variabilità della frequenza cardiaca (HRV) e, nello specifico, il parametro *RMSSD (Root Mean Square of Successive Differences)*. In altre parole, se l'EDA fa da indicatore per l'aumento dello stress, la variabilità cardiaca fa vedere se il 'freno vagale' sta ancora resistendo. Un crollo improvviso di questa variabilità, unito a un'alta sudorazione, è indice di una informazione delicata e precisa: in quel momento il soggetto ha esaurito del tutto le sue difese interne per gestire il sovraccarico emotivo.

4.8 Il protocollo clinico e la definizione delle classi (*labeling*)

La gestione pratica delle sessioni all'interno della stanza multisensoriale e la successiva compilazione del diario clinico sono state condotte in prima linea dagli operatori sanitari ed educativi di riferimento dei soggetti. Il presente lavoro bioingegneristico si è collocato nella fase immediatamente successiva, affrontando la complessa sfida della conversione di appunti testuali, orari e descrizioni qualitative in un formato digitale, etichettato e leggibile dall'algoritmo di *Machine Learning*.

Per evitare assegnazioni arbitrarie, la categorizzazione degli eventi ha trovato solide fondamenta nel "modello circomplesso dell'affetto" di Russell (1980) [28]. Questo costruito modella gli stati psicofisiologici come coordinate all'interno di uno spazio bi-dimensionale definito da due assi: l'*arousal* (attivazione fisica e intensità) e la *valenza* (piacevolezza o spiacevolezza dell'esperienza). Ispirandosi a questo modello teorico e osservando le metodologie di validazione di dataset accademici di riferimento (come il *WESAD - wearable stress and affect detection*), le registrazioni sono state frammentate in tre macro-classi comportamentali:

- **Classe 0 (Baseline / Relax):** Rappresenta lo stato di calma fisiologica (basso livello di *arousal*, valenza neutra o lievemente positiva). In questa condizione il sistema nervoso non è sottoposto ad alcuno sforzo allostatico ed è dominato dall'azione del tono parasimpatico. Questo si traduce in un'attività elettrodermica piatta (assenza di picchi SCR) e in una spiccata variabilità cardiaca (RMSSD elevato). Nel protocollo, l'etichetta '0' è stata assegnata ai momenti di quiete o di rilassamento profondo, come quello indotto dal letto vibroacustico. Isolare questi segmenti ha fornito all'algoritmo il "punto zero" fisiologico di ogni soggetto, parametro indispensabile per consentire alla *Change Point Analysis* di intercettare i successivi sbalzi di stress.
- **Classe 1 (Carico cognitivo/ Engagement):** identifica uno stato di attivazione finalizzata e consapevole (*arousal* medio-alto, valenza decisamente positiva). Rientrano in questa fondamentale categoria le fasi temporali in cui i soggetti

erano fortemente coinvolti in compiti attivi, come le interazioni con i tubi luminosi o i lavori di stimolazione somatica. Questa classe permette di insegnare al modello a non confondere il *pattern* del sano coinvolgimento (ad es. alta frequenza di picchi EDA, ma con ampiezza moderata) con l'imminenza di una crisi, generando falsi allarmi.

- **Classe 2 (Stress / Meltdown):** Rappresenta l'apice del sovraccarico emotivo e sensoriale (*arousal* estremo, valenza profondamente negativa). Fisiologicamente si manifesta con una violenta attivazione del sistema simpatico, caratterizzata da picchi EDA altissimi e ravvicinati a fronte di un crollo delle difese vagali. Tale condizione si è verificata nella sessione del 16 novembre, quando il Soggetto 3 ha vissuto un forte attacco di panico nei primi minuti di acquisizione. L'intervento medico tempestivo ha permesso di campionare, nell'ora successiva, l'intera fase di recupero verso la normalità. Sebbene l'isolamento di questi minuti sia stato prezioso per estrarre l'autentica prova biologica del *meltdown*, il campione è risultato quantitativamente insufficiente per l'addestramento di un modello di *Machine Learning*.

La sproporzione tra le ore di quiete e i pochi minuti di crisi reale ha determinato un severo sbilanciamento delle classi. Questa limitazione ha rappresentato la svolta metodologica del lavoro: per fornire al modello un volume di dati sufficiente a riconoscere i segnali di pericolo, si è reso indispensabile ricorrere alla *Time Series Data Augmentation* e alla generazione di pazienti virtuali.

5. Pre-processing dei segnali

5.1 Dal diario clinico al dato digitale: la definizione del ground truth

Il primo problema che si è dovuto affrontare iniziando a lavorare sull'architettura software è stato tradurre la realtà dell'ambiente clinico in un formato logico e matematico comprensibile a MATLAB. Addestrare un algoritmo predittivo con un flusso continuo di dati grezzi, comprensivo di artefatti da movimento o di pause irrilevanti, sarebbe stato un grave errore metodologico.

Per questo motivo, si è deciso di organizzare il lavoro in blocchi indipendenti. Il primo script è dedicato esclusivamente alla definizione dell'esperimento e alla generazione del *ground truth* (o verità di base) che, in ambito di *Machine Learning*, rappresenta il dato certo su cui il modello dovrà basarsi per imparare a fare previsioni. Prima di elaborare i segnali, è stato necessario isolare le esatte finestre temporali di stress o relax, basandosi su quanto documentato dal protocollo clinico.

Innanzitutto, in questa fase è stata cruciale la gestione informatica dei file, poiché, avendo a che fare con enormi moli di dati, divisi per giorno di acquisizione, si è evitata la programmazione di percorsi rigidi che avrebbero reso il codice inutilizzabile su altri terminali. Dunque, è stata implementata un'astrazione del *file system* definendo la directory principale tramite la singola variabile (`base_path`) utilizzando la funzione nativa di MATLAB `fullfile` per ricostruire dinamicamente i percorsi. A questo è stata affiancata una logica di controllo di sicurezza: prima di avviare qualsiasi caricamento, la funzione `exist` verifica che le cartelle dei dati siano fisicamente presenti, bloccando l'esecuzione con un messaggio di errore in caso di anomalie e rendendo il codice robusto.

Risolta l'infrastruttura di lettura, l'obiettivo è stato la digitalizzazione degli appunti qualitativi del diario clinico redatto dagli operatori(es. "Soggetto 3, 16 Nov, ore 17:47, fase di recupero post-crisi"). Queste descrizioni umane sono state trasformate in un *array* di strutture dati MATLAB, denominato `events`. A ogni evento è stato associato un identificativo univoco (`id`), il nome del paziente (`soggetto`) e il puntatore alla

cartella dei suoi dati grezzi (*folder*).

Nella definizione di questa struttura si è dovuto affrontare l'insidia tecnica più critica di tutta la fase di acquisizione: il disallineamento dei fusi orari. Il braccialetto Empatica E4, come molti *wearable* medicali, registra i *timestamp* basandosi sul tempo coordinato universale (UTC). Tuttavia, gli appunti sul diario clinico erano stati logicamente presi seguendo l'orario locale italiano, ma ignorare questa discrepanza hardware-software avrebbe compromesso irreversibilmente l'addestramento, allineando matematicamente un picco di panico con l'etichetta comportamentale di un momento di relax. Per ovviare al problema, sono state fissate le coordinate di inizio (*start*) e fine (*end*) di ogni evento utilizzando la funzione `datetime`, forzando esplicitamente l'allineamento temporale tramite il parametro `'TimeZone'`, `'Europe/Rome'`. Questo passaggio ha garantito una sovrapposizione perfetta, al millisecondo, tra il dato fisiologico e l'annotazione clinica.

Contestualmente alla delimitazione della finestra temporale, si è compilato il campo `label`, assegnando a ogni evento il proprio valore *target* (0, 1 o 2), corrispondente alle macro-classi comportamentali definite nel capitolo precedente. In questo modo si è istruito il modello sul reale significato psicofisiologico dei segmenti in analisi (*data labeling*).

Infine, per mantenere l'architettura modulare, queste informazioni sono state rese persistenti salvando la struttura `events` nel file fisico `protocollo.mat`. Questo file rappresenta l'anello di congiunzione tra l'esperimento clinico e l'elaborazione matematica: il modulo di *processing* successivo dovrà semplicemente caricare questo file e processare le finestre temporali che vi sono descritte, ottenendo una *pipeline* fluida e scientificamente riproducibile.

5.2 Lettura dei file sorgente e sincronizzazione temporale assoluta

Il fulcro dell'architettura di analisi risiede nella fase di *processing*. La funzione primaria di questo blocco è l'acquisizione dei dati grezzi estratti dall' E4, per trasformarli in matrici pronte per l'analisi. Tuttavia, la fase di caricamento dei dati si è scontrata

con un problema strettamente legato alla codifica dei file in formato CSV. Probabilmente a causa di un precedente passaggio su sistemi operativi Windows, i documenti presentavano all'inizio dei caratteri nascosti chiamati BOM (*Byte Order Mark*). Pur risultando invisibili nei comuni fogli di calcolo, per l'ambiente MATLAB rappresentavano un ostacolo critico: il software interpretava il primo valore del documento (il *timestamp* di avvio) non come un numero, ma come una stringa di testo. Questa errata decodifica generava un errore a catena di tipo NaT (*Not a Time*), facendo perdere all'algoritmo il suo "punto zero" e compromettendo l'intero asse temporale. Per risolvere questa criticità senza ricorrere a complessi cicli di pulizia testuale, si è utilizzata l'istruzione MATLAB `readmatrix(..., 'OutputType', 'double')`. Questo comando impone al software di ignorare in automatico qualsiasi carattere invisibile o lettera, estraendo esclusivamente i dati numerici e garantendo un'importazione sicura.

Risolto l'errore di lettura, si è posto il problema di come interpretare correttamente l'enorme mole di numeri appena importata. I file generati dal braccialetto Empatica E4 sono infatti estremamente essenziali: per ottimizzare lo spazio di memoria, non contengono etichette di testo che spieghino il contenuto delle colonne (come ad esempio le parole "Data", "Ora" o "Valore"). Al contrario, i dati sono organizzati secondo una struttura fissa e predeterminata dal produttore. Per questo motivo, si è dovuto "insegnare" all'algoritmo a leggere il documento basandosi esclusivamente sulla posizione delle singole righe.

L'estrazione dei parametri fondamentali è avvenuta quindi in tre passaggi mirati: dalla cella in riga 1 è stato prelevato il *time zero* (il *timestamp* Unix di avvio sensore); dalla riga 2 è stata estratta la frequenza di campionamento (f_s); infine, dalla terza riga in poi, l'algoritmo ha isolato esclusivamente la matrice storica dei biosegnali.

Superata l'estrazione di tali parametri, il processo di pre-elaborazione ha previsto un fondamentale intervento di correzione dimensionale, fortemente necessario quando si integrano sensori caratterizzati da nature fisiche e frequenze di campionamento profondamente differenti (ad esempio, l'EDA viene campionata a 4 Hz, mentre l'accelerometro a 32 Hz). Un biosegnale come l'attività elettrodermica (EDA) è unidimensionale e popola un'unica colonna della matrice; al contrario, l'accelerometro triassiale (ACC) rileva il movimento nello spazio generando tre valori simultanei (sui vettori spaziali

X, Y e Z) per ogni singolo istante temporale. Dinanzi a questa discrepanza strutturale, se il codice si fosse limitato a contare alla cieca tutte le celle piene di numeri presenti nel file, si sarebbe verificato un grave artefatto matematico: il segnale del movimento, avendo tre colonne di valori, sarebbe stato interpretato come se fosse lungo esattamente il triplo rispetto al segnale EDA. Sfruttando il comando MATLAB `size(data, 1)`, il sistema è stato forzato a contare unicamente il numero di campioni (le righe), ignorando la tridimensionalità delle colonne e garantendo la perfetta sovrapposibilità matematica tra segnali di diversa natura fisica.

L'ultimo ed importante tassello ha riguardato la ricostruzione dell'asse temporale. Per risparmiare memoria, i file grezzi non stampano l'orario completo accanto a ogni misurazione. Per associare una variazione fisiologica all'esatto orario del diario clinico, si è reso necessario ricreare matematicamente l'orario di ogni singolo campione applicando in modo sequenziale la seguente relazione:

$$t(i) = T_{start} + \frac{i - 1}{f_s} \quad (5.1)$$

Il termine T_{start} rappresenta il *timestamp* iniziale di accensione del dispositivo (forzato nel fuso orario di Roma), i indica l'indice del campione (numero della riga corrente), mentre f_s rappresenta la frequenza di campionamento del sensore.

Dividendo l'indice del campione per la frequenza (ad esempio 4 Hz per l'EDA), l'algoritmo calcola esattamente lo scarto di tempo trascorso dall'inizio della registrazione, generando un asse temporale vettoriale con risoluzione al millisecondo (es. 0.25 s, 0.50 s, 0.75 s).

Il risultato di questo lavoro di pulizia e ricostruzione matematica è una base di partenza solida, in cui le fluttuazioni biologiche invisibili si sovrappongono esattamente agli eventi clinici annotati dai terapisti.

5.3 Estrazione e processamento delle metriche cardiovascolari (HRV)

Come analizzato nel capitolo precedente, la variabilità della frequenza cardiaca (HRV) costituisce un parametro clinico importante per quantificare l'efficienza del "freno vagale" e la capacità autoregolativa di un soggetto con ASD. L'estrazione di questa informazione non si basa sulla semplice frequenza cardiaca media (HR), che appiattirebbe le fluttuazioni microscopiche necessarie per l'analisi autonoma, ma sull'elaborazione della sequenza esatta degli intervalli IBI (*Inter-Beat Interval*), esportata dal dispositivo nel file `IBI.csv`.

Parallelamente alla gestione dei segnali continui (come EDA e accelerometro), l'architettura software ha richiesto lo sviluppo di un modulo specifico per l'elaborazione del tracciato cardiaco, realizzato in stretta sinergia con la collega co-autrice del progetto.

La criticità intrinseca del segnale IBI risiede nella sua natura asincrona. Infatti, a differenza della conduttanza cutanea o del movimento, che vengono campionati dal sensore a un ritmo fisso e costante (ad esempio 4 o 32 volte al secondo), i battiti del cuore non obbediscono a una frequenza di campionamento fissa, poiché la distanza temporale tra un battito e l'altro cambia continuamente. Per gestire questa irregolarità vettoriale, è stata strutturata una funzione *ad hoc* denominata `load_empativa_ibi`, il cui compito non è semplicemente il caricamento numerico, ma la precisa associazione di ogni singolo intervallo R-R (la distanza tra due picchi cardiaci) al suo esatto e assoluto istante temporale, sincronizzandolo perfettamente con il resto dei dati clinici.

Una volta importati i dati grezzi, il focus dell'elaborazione si sposta sulla funzione `calc_hrv`, progettata per estrarre le metriche fondamentali dell'HRV nel dominio del tempo. Tuttavia, prima di eseguire qualsiasi formula matematica, il codice implementa un rigoroso sistema di sicurezza statistica, un passaggio fondamentale quando si lavora con soggetti con ASD. Durante un episodio di stress o *stimming*, i bruschi movimenti del polso causano inevitabili distacchi temporanei del sensore ottico PPG, generando lacune nell'acquisizione.

Dunque, per evitare il rischio di elaborare dati corrotti o incompleti, la funzione analizza ogni singola finestra temporale ed effettua un rigoroso conteggio dei battiti validi registrati dal sensore. In questa fase è stata implementata una stringente regola di validità: se all'interno della finestra in esame il dispositivo ha rilevato meno di 5 battiti consecutivi, il frammento viene etichettato come inaffidabile. La scelta di questa soglia minima non è casuale, ma risponde a una precisa esigenza fisiologica e ingegneristica; tentare di calcolare indici di variabilità complessi (come una deviazione standard) partendo da un campione di sole tre o quattro misurazioni ($N < 5$) produrrebbe un risultato matematicamente calcolabile, ma clinicamente del tutto privo di significato. Il numero restituito sarebbe casuale e incapace di riflettere la reale modulazione del sistema nervoso autonomo.

Di fronte a una simile frammentazione del segnale, l'algoritmo interviene in modo conservativo: scarta in automatico la finestra temporale compromessa e ne forza il risultato a un valore nullo, identificato in ambiente MATLAB dall'etichetta NaN (*Not a Number*). Questa rigorosa scelta progettuale agisce come un vero e proprio filtro di sicurezza, impedendo che il successivo modello di *Machine Learning* venga addestrato su dati "inquinati", portandolo ad associare falsi *pattern* di stress a puri artefatti motori.

Superato con successo questo controllo di validità, la funzione estrae il parametro principale per la valutazione del rilassamento e dello stress acuto: l'RMSSD (*Root Mean Square of Successive Differences*). Questa metrica non valuta il battito cardiaco in sé, ma la variazione istantanea tra un battito e l'altro, riflettendo così in modo diretto l'azione frenante e calmante del nervo vago.

Invece di affidarsi a librerie preimpostate, il calcolo è stato implementato passo dopo passo per mantenere il totale controllo logico sui dati. La matematica dietro l'RMSSD opera secondo la seguente formula:

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (IBI_{i+1} - IBI_i)^2} \quad (5.2)$$

Inizialmente, l'algoritmo isola due battiti cardiaci consecutivi e calcola l'esatta differenza temporale che intercorre tra il primo e il secondo intervallo. Questa sottrazione sequenziale viene ripetuta per tutte le coppie di battiti presenti nella finestra temporale

in esame. Successivamente, ogni singola differenza ottenuta viene elevata al quadrato. Questo specifico passaggio riveste un'importanza clinica cruciale poiché l'elevamento a potenza agisce matematicamente come un filtro amplificatore: da un lato smussa le variazioni fisiologiche più lievi e costanti (come la naturale aritmia sinusale indotta dalla respirazione), dall'altro dà il massimo risalto ai cambiamenti bruschi e improvvisi. Un salto repentino nella durata tra due battiti è infatti il segnale inequivocabile di una rapida reazione nervosa. La media aritmetica di questi quadrati viene poi riportata alla dimensione temporale originaria tramite l'estrazione della radice quadrata, restituendo il dato nell'unità di misura clinica standard (millisecondi, ms).

In parallelo, dalla medesima matrice di dati validati, la funzione estrae un parametro complementare: la SDNN (*Standard Deviation of Normal-to-Normal intervals*). La necessità di affiancare l'RMSSD a questo secondo indice deriva dai loro diversi significati fisiologici. Mentre l'RMSSD si concentra sugli scarti tra battiti adiacenti (catturando la risposta veloce e fasica del sistema parasimpatico), la deviazione standard globale della SDNN valuta quanto l'intero *set* di battiti della finestra si discosti dal ritmo medio, riflettendo la stabilità e la naturale capacità di adattamento a lungo termine del sistema.

Questa estrazione congiunta fornisce al futuro algoritmo predittivo una visione autonoma completa, in grado di mappare sia la reattività istantanea del paziente a uno stimolo acuto, sia il suo equilibrio nervoso di fondo.

5.4 Decomposizione dell'attività elettrodermica tramite ottimizzazione convessa

Come ampiamente discusso nei capitoli precedenti, l'attività elettrodermica (EDA) costituisce un indicatore primario per oggettivare lo stato di *arousal* del paziente ASD. Tuttavia, il tracciato grezzo acquisito dal braccialetto Empatica E4 si presenta visivamente come una singola linea continua che, dal punto di vista biologico, nasconde in realtà la sovrapposizione di due fenomeni distinti: la componente tonica (SCL, *Skin Conductance Level*), lenta e legata all'idratazione generale della pelle, e la componen-

te fasica (SCR, *Skin Conductance Response*), velocissima e rappresentativa delle reali scariche nervose simpatiche di allerta [30].

Se si fosse scelto di addestrare l'algoritmo di *Machine Learning* fornendogli in ingresso il segnale grezzo, il modello non avrebbe avuto gli strumenti per distinguere una lenta e normale sudorazione termica da un picco improvviso causato da un sovraccarico emotivo. Per questo motivo, separare in modo netto queste due componenti non è stata solo un'operazione estetica per "pulire" i grafici, ma un passaggio indispensabile per dare senso clinico ai dati.

L'approccio metodologico più comune per la pulizia della conduttanza cutanea è quello di utilizzare dei classici filtri passa-banda (come i filtri gaussiani o Butterworth). Tuttavia, questi strumenti, seppur tradizionali si limitano a smussare le frequenze del segnale per tentativi, alterando spesso la forma originale del dato clinico facendo perdere la sua utilità informativa ai fini dell'analisi. Per una elaborazione di qualità da un punto di vista analitico, si è scelto di non adottare il filtraggio classico a favore di un metodo che vede in primo piano l'algoritmo cvxEDA, proposto e validato in letteratura da Greco et al. (2014) [25].

L'innovatività di questo approccio risiede nel fatto che l'algoritmo non tratta l'EDA come una semplice onda da smussare, ma affronta la questione come un vero e proprio problema inverso: a partire dall'attività elettrodermica misurata sulla superficie del polso (l'effetto), il sistema sfrutta la matematica per ricostruire a ritroso la complessa e invisibile attività del sistema nervoso autonomo (la causa). Per fare questo, il segnale EDA grezzo viene modellato non come un'entità unica, ma come la somma lineare di tre componenti ben distinte:

$$y(t) = r(t) + t(t) + \epsilon(t) \quad (5.3)$$

In questa equazione fondamentale, $y(t)$ rappresenta il segnale misurato dal sensore Empatica, $r(t)$ è la risposta fasica (veloce), $t(t)$ è il livello tonico (lento) ed $\epsilon(t)$ rappresenta il rumore di misura. Poiché da un singolo segnale $y(t)$ è impossibile ricavare matematicamente le altre due incognite senza ulteriori regole, il codice implementato traduce la biologia umana in vincoli matematici, operando su tre pilastri.

1. Il Modello fisiologico della ghiandola (la componente fasica): il primo passo è quello di isolare la componente fasica $r(t)$, che rappresenta la risposta rapida e istantanea del soggetto a uno stimolo o a un picco di stress. Da un punto di vista puramente fisiologico, quando il cervello invia una scarica di impulsi nervosi (il *driver* neurale), le ghiandole non hanno una risposta assimilabile a un gradino perfetto che si esaurisce al cessare dello stimolo; al contrario, essa risulta alquanto asimmetrica, poiché si riempiono di sudore in modo quasi istantaneo, ma impiegano molto più tempo per svuotarsi o far evaporare il liquido. Questa meccanica dei fluidi viene descritta nel codice attraverso la funzione bi-esponenziale di Bateman:

$$b(t) = e^{-t/\tau_0} - e^{-t/\tau_1} \quad (5.4)$$

Questa formula disegna l'esatta forma dell'impulso della ghiandola. I due parametri temporali sono stati impostati secondo i rigorosi standard di letteratura: un tempo di salita rapidissimo ($\tau_1 = 0.7$ s) e un tempo di decadimento molto più lento ($\tau_0 = 3.0$ s). Di conseguenza, il segnale fasico totale r che si osserva non è altro che il risultato matematico della convoluzione tra i veri impulsi nervosi (definiti dal vettore p) e questa forma fisica della ghiandola (rappresentata da una matrice sparsa A). In sintesi, $r = A \cdot p$. L'obiettivo ultimo di questa complessa operazione matematica, e il vero valore aggiunto a livello clinico del presente lavoro consiste nell'estrarre il vettore p . Questo parametro, infatti, rappresenta l'autentica e nitida attività nervosa sudomotoria: estrarlo significa poter isolare i singoli impulsi del sistema simpatico nella loro reale tempestività, depurandoli completamente dal ritardo e dalla lenta inerzia biomeccanica della ghiandola sudoripara.

2. La curva di base (La componente tonica): Il secondo elemento da isolare è la componente tonica $t(t)$, ovvero il livello di base della conduttanza cutanea (SCL). Questo valore riflette lo stato di idratazione generale della pelle e varia in modo estremamente lento e progressivo durante l'intera seduta (la cosiddetta deriva o *drift*). Per modellarla in modo realistico, evitando l'errore di appiattare i dati con una banale media mobile, l'algoritmo implementato [25] utilizza un approccio molto più elegante basato sulle *cubic b-splines*. Dunque, invece di ostinarsi a calcolare un valore per ogni singolo mil-

lisecondo registrato, il codice semplifica il problema definendo dei “punti di controllo”, chiamati nodi, posizionati a intervalli temporali regolari (nello specifico, uno ogni 10 secondi). L’obiettivo dell’algoritmo è far passare attraverso questi nodi una curva matematica morbidissima, capace di seguire l’andamento lento del sudore senza presentare salti innaturali. L’equazione che governa questa componente diventa quindi:

$$t = B \cdot l + C \cdot d \quad (5.5)$$

Ognuno dei termini che compongono la formula hanno una funzionalità ed un significato geometrico ben preciso; nello specifico, essi operano in questo modo:

- **La matrice B:** Questa matrice non contiene i dati fisiologici del soggetto, ma rappresenta l’infrastruttura matematica universale delle *cubic b-splines*. Funge da vero e proprio “scheletro” poiché contiene le basi polinomiali di terzo grado, ovvero le regole geometriche fisse che garantiscono che la curva generata sia sempre perfettamente fluida, continua e priva di spigoli (garantendo la cosiddetta continuità C^2).
- **Il vettore l (le altezze dei nodi):** rappresenta l’incognita fondamentale che l’algoritmo deve scoprire. È un vettore che contiene i coefficienti della spline, che visivamente si possono immaginare come le “altezze” o i “pesi” di ciascun nodo. Modificando i valori di l , il risolutore matematico “tira” la curva verso l’alto o verso il basso nei vari istanti temporali, fino a farla combaciare in modo ottimale con la reale sudorazione lenta del paziente. La moltiplicazione $B \cdot l$ genera così l’onda morbida della conduttanza.
- **Il termine $C \cdot d$ (la deriva lineare):** Anche la pelle di un soggetto perfettamente a riposo possiede una sua naturale pendenza di fondo (ad esempio, il soggetto potrebbe sudare in modo lievissimo ma costante a causa della temperatura della stanza, oppure rilassarsi progressivamente perdendo conduttanza). Per catturare questo andamento globale, il sistema affianca alle curve spline una retta inclinata. La matrice fissa C fornisce le “coordinate” del tempo, mentre il vettore incognito d contiene i due moltiplicatori essenziali: l’intercetta (il punto di partenza iniziale) e la pendenza (il *trend* o *drift* lineare della seduta).

Sommare le morbide onde delle spline ($B \cdot l$) alla retta di fondo ($C \cdot d$) permette di ricostruire un tracciato tonico $t(t)$ fisiologicamente autentico e realistico, pronto per essere sottratto dal segnale globale e lasciare emergere la pura componente nervosa.

3. L'Ottimizzazione convessa: A questo punto, avendo definito la parte teorica, l'algoritmo deve materialmente scoprire i valori esatti degli impulsi nervosi (ovvero il vettore p) e le altezze della componente lenta (il vettore l). Per farlo, il codice non applica una formula algebrica diretta, poiché il problema inverso ammette infinite soluzioni possibili, ma utilizza un sofisticato algoritmo di ottimizzazione iterativa, sfruttando il risolutore quadprog (programmazione quadratica) nativo dell'ambiente MATLAB. L'obiettivo del *solver* è esplorare le combinazioni di valori fino a trovare quella che minimizza una complessa equazione chiamata funzione di costo (o *Loss Function*, indicata con J). L'equazione è la seguente:

$$J = \frac{1}{2} \|y - (Ap + Bl + Cd)\|_2^2 + \alpha \|p\|_1 + \frac{\gamma}{2} \|l\|_2^2 \quad (5.6)$$

Questa formula costringe l'algoritmo a trovare un delicato compromesso tra tre diverse esigenze biologiche e matematiche:

- **L'errore di ricostruzione:** Il primo termine della formula valuta la distanza (al quadrato, per penalizzare severamente gli errori grandi) tra il segnale originale misurato dal braccialetto (y) e il segnale artificiale che il software sta tentando di ricostruire unendo fasica, tonica e deriva. L'algoritmo è obbligato a mantenere questa differenza il più vicina possibile allo zero, garantendo che la simulazione matematica non si discosti dalla realtà clinica osservata.
- **La sparsità neurale (Il parametro α):** Il secondo termine introduce una penalizzazione matematica basata sulla norma L1 (la somma dei valori assoluti degli impulsi). Fisiologicamente, il sistema nervoso simpatico non "spara" segnali in continuazione, ma invia impulsi rari, netti e ben distanziati (un comportamento, appunto, "sparso"). L'introduzione del parametro α in questo passaggio funge da severo criterio di selettività, poiché obbliga matematicamente il risolutore a ignorare le deboli e continue fluttuazioni di base. In questo modo, l'algoritmo

azzera del tutto il rumore e isola esclusivamente i veri picchi di attivazione del sistema simpatico, restituendoli in modo netto.

- **La lisciatura (Il parametro γ):** Il terzo termine dell'equazione utilizza il parametro γ per applicare una penalizzazione matematica (basata sulla norma L2, ovvero la somma dei quadrati) direttamente ai coefficienti l della componente tonica. Lo scopo di questo vincolo è allineare il calcolo alla reale fisiologia del corpo umano: poiché il livello tonico riflette l'idratazione di base della pelle, esso si modifica per natura in modo estremamente lento. La penalizzazione impedisce quindi alla curva di base di registrare salti improvvisi. Di conseguenza, quando l'algoritmo rileva una variazione rapida nel segnale originario, è costretto ad assegnarla esclusivamente all'attività del sistema nervoso (la componente fasica), garantendo che la componente tonica mantenga sempre un andamento fluido, graduale e privo di sbalzi innaturali.

5.4.1 Il vincolo di non-negatività e la correzione di offset

All'algoritmo viene imposta una rigorosa regola biologica di partenza: i valori degli impulsi nervosi calcolati devono essere strettamente non negativi ($p \geq 0$). Questo avviene perché, da un punto di vista puramente fisiologico, il sistema nervoso simpatico svolge unicamente un'azione eccitatoria sulle ghiandole sudoripare, stimolando la produzione di sudore. Il corpo umano non possiede, infatti, un meccanismo neurale capace di inviare un "impulso negativo" per riassorbire attivamente e istantaneamente il liquido secreto sulla pelle.

L'introduzione di questo fondamentale vincolo matematico è necessaria per spiegare una peculiare anomalia riscontrata visivamente nei tracciati di alcuni soggetti. Questo fenomeno si è manifestato in soggetti caratterizzati da livelli di conduttanza di base estremamente bassi (una condizione nota in letteratura come *low SCL* o *non-responders*, spesso legata a una pelle molto secca o a un contatto debole con il sensore).

Analizzando questi specifici casi, si è notato un comportamento ricorrente: quando il segnale grezzo era prossimo allo zero (ad esempio, $0.1 \mu\text{S}$) e l'algoritmo rilevava

un'improvvisa attivazione nervosa stimando un picco fasico più alto (ad esempio, $0.5 \mu\text{S}$), il risolutore matematico andava incontro a un vero e proprio limite strutturale. Per mantenere valido l'equilibrio dell'equazione primaria della decomposizione ($y = t + r$, ovvero $0.1 \mu\text{S} = t + 0.5 \mu\text{S}$), il sistema era matematicamente costretto a compensare la differenza, spingendo la componente tonica verso valori negativi. Calcolando quindi un valore di t pari a $-0.4 \mu\text{S}$, pur di far spazio a un picco fasico evidente su una linea di base quasi nulla, l'algoritmo generava un abbassamento innaturale della componente lenta.

Aver compreso passo dopo passo la matematica che sta dietro a questo artefatto ha permesso di affrontarlo e risolverlo senza dover scartare le acquisizioni o invalidare l'analisi del soggetto. Per restituire piena coerenza clinica alle metriche e garantire che il successivo modello predittivo ricevesse in ingresso esclusivamente valori positivi e realistici, si è scelto di intervenire nella fase di post-decomposizione applicando una mirata correzione di *offset*.

A livello di codice, questa operazione è stata implementata in modo molto lineare: si individua il punto di minimo assoluto raggiunto dalla curva tonica appena generata e, se questo valore scende sotto lo zero, l'algoritmo ne calcola il valore assoluto e lo somma uniformemente a tutto il tracciato. Attraverso questa semplice traslazione verticale, l'intera componente tonica viene riportata sopra la soglia dello zero.

Ai fini dell'analisi predittiva, infatti, il valore numerico assoluto della conduttanza è trascurabile rispetto alla sua dinamica: ciò che conta realmente per valutare un sovraccarico emotivo è la pendenza della curva (il cosiddetto *slope*, ampiamente validato in letteratura) e l'ampiezza delle sue variazioni nel tempo. Traslare il grafico verticalmente di pochi centesimi di microSiemens risolve l'incongruenza biologica, preservando al contempo la complessa morfologia del tracciato e i tempi di discesa estratti dall'ottimizzazione convessa. Inoltre, questo intervento non inficia in alcun modo la peculiare componente fasica r , che rimane pura e fedele all'estrazione originale.

5.5 Valutazione motoria e continuità del monitoraggio

Per completare il processo di elaborazione dei segnali, non ci si è limitati ad analizzare solo l'attivazione autonoma, ma è stata affiancata ad essa un'analisi cinematica di fondamentale importanza, processando i dati provenienti dall'accelerometro triassiale del dispositivo *wearable*.

L'accelerometro registra le accelerazioni del polso del paziente lungo tre assi spaziali ortogonali (X , Y e Z). Tuttavia, si deve tenere conto che utilizzare questi tre assi separatamente comporterebbe un grave limite pratico e conseguentemente i valori misurati cambierebbero drasticamente a seconda di come il paziente indossa il bracciale o di come ruota il braccio nello spazio (ad esempio, un movimento orizzontale potrebbe essere registrato sull'asse X in un momento, e sull'asse Y un secondo dopo se il polso viene girato). Per svincolarsi completamente dalla posizione spaziale e dall'orientamento del sensore, il codice esegue un'operazione matematica di compressione calcolando la Magnitudo Vettoriale (nota anche come norma Euclidea o VMU).

Grazie a questo calcolo, i tre vettori dell'accelerometro vengono riassunti in una sola grandezza matematica. Il risultato è una misura pura che indica quanto il paziente si stia muovendo, svincolando l'analisi dalla specifica direzione del movimento. Nel contesto specifico del Disturbo dello Spettro Autistico (ASD), questo dato assume un'importanza clinica inestimabile; l'autismo è infatti spesso caratterizzato da fenomeni di irrequietezza motoria o da movimenti ripetitivi e stereotipati (il cosiddetto *stimming*, come lo sfarfallio delle mani). Per questo motivo, avere a disposizione la VMU fornisce al successivo modello predittivo un contesto comportamentale rilevante che permette all'algoritmo di capire se un improvviso picco di conduttanza cutanea è dovuto a un reale sovraccarico emotivo interno, oppure se è semplicemente il riflesso fisiologico di un'intensa attività motoria in corso.

Infine, pensando a un utilizzo del sistema nella realtà clinica di tutti i giorni, si è rivelato indispensabile progettare una vera e propria procedura di sicurezza (nota come algoritmo di *fallback*).

L'algoritmo cvxEDA è uno strumento matematico tanto raffinato quanto delicato. Nella pratica clinica reale, specialmente quando si monitorano soggetti che possono compiere movimenti bruschi o inaspettati, capita di frequente che il sensore venga urtato. Questi eventi generano forti artefatti da movimento che “sporcano” il segnale elettrodermico. Di fronte a un tracciato così disturbato, le equazioni alla base del problema inverso non riescono più a trovare una soluzione logica; di conseguenza, il calcolo matematico va in errore e rischia di far bloccare l'intero algoritmo.

Per evitare che il sistema vada in *crash* proprio durante il monitoraggio di un paziente, è stata inserita una regola di controllo: se l'ottimizzazione matematica si rende conto di non poter processare il segnale, il software intercetta l'errore senza spegnersi e passa in automatico a un “piano B”.

In questo scenario di emergenza, l'estrazione della componente lenta e la pulizia del segnale vengono temporaneamente affidate a filtri tradizionali molto più robusti, come il passa-basso Butterworth o la lisciatura gaussiana. Sicuramente si tratta di un compromesso: passando a questi filtri si perde temporaneamente la precisione chirurgica con cui il cvxEDA isola i singoli impulsi nervosi. Tuttavia, dal punto di vista clinico, questa scelta pratica assicura che il monitoraggio non si interrompa mai e continui a fornire al modello predittivo un dato di base pulito su cui lavorare. In questo modo, garantiamo che l'osservazione e la tutela del paziente restino sempre attive, persino nelle condizioni di acquisizione peggiori.

5.6 Analisi multisegnale e risultati del pre-processing

Definita e implementata l'architettura software per la pulizia, la sincronizzazione e la scomposizione dei segnali, come descritto ampiamente nelle sezioni precedenti, prima di procedere all'estrazione delle *feature* e al successivo addestramento del modello di *Machine Learning*, è opportuno verificare che i dati elaborati dall'algoritmo non siano solamente corretti da un punto di vista matematico, ma che possiedano un coerente e reale significato clinico. L'obiettivo di questa sezione è quello di validare i risultati del

pre-processing, dimostrando come l'algoritmo sviluppato sia stato in grado di catturare e tradurre in modo fedele le reali reazioni fisiologiche dei pazienti.

I tracciati in esame riflettono le dinamiche delle due sessioni di lavoro (16 e 18 novembre). Come ampiamente descritto nel capitolo precedente relativo al protocollo clinico, le acquisizioni si sono svolte all'interno della stanza multisensoriale *Snoezelen*, esponendo i soggetti a un ambiente altamente modulabile caratterizzato da forti contrasti visivi (luce di Wood, fibre ottiche) e stimolazioni propriocettive (materasso vibroacustico, pressione profonda). La prima sessione è stata caratterizzata da un approccio più libero, focalizzato sull'esplorazione sensoriale spontanea e sulla gestione di stati emotivi pregressi. La seconda sessione, invece, ha previsto un intervento clinico più strutturato, basato sull'uso di forti contrasti visivi e sul contatto fisico diretto.

Come risultato finale della *pipeline* informatica, per ogni singolo evento annotato sul protocollo clinico, il sistema ha generato una *dashboard* grafica composta da cinque riquadri sovrapposti e sincronizzati al millisecondo. Questo strumento visivo allinea l'attività elettrodermica globale, la sua scomposizione in componente tonica e fasica (ottenute tramite ottimizzazione convessa), l'intensità del movimento (magnitudo dell'accelerometro) e i parametri cardiovascolari (frequenza cardiaca e variabilità HRV).

Incrociando l'andamento di queste *dashboard* con le annotazioni comportamentali prese minuto per minuto nel protocollo, è possibile osservare la perfetta coerenza tra il dato biologico estratto dalla matematica e il comportamento reale osservato nella stanza. Di seguito vengono dunque presentati i risultati relativi all'analisi di tre soggetti, esaminati nel dettaglio per dimostrare la solidità clinica e l'efficacia dell'infrastruttura di analisi finora costruita.

5.6.1 Il caso del Soggetto 1

Il caso del Soggetto 1 rappresenta un esempio clinico efficace per comprendere l'impatto fisiologico della stimolazione multisensoriale e, in particolare, della terapia vibroacustica. L'analisi dei dati raccolti permette di osservare come il sistema nervoso autonomo reagisca positivamente agli stati di rilassamento profondo indotti meccanicamente e come questa ritrovata calma influenzi in modo diretto la gestione dei momenti

di transizione e di interazione sociale. La prima sessione di analisi mostrerà il passaggio dalla ricezione sensoriale passiva al gioco attivo, mentre la seconda consentirà di valutare la stabilità di questo equilibrio fisiologico nel tempo e sotto forte stimolazione meccanica.

Sessione 1: Analisi della sessione sensoriale

Durante questa sessione, il Soggetto 1 ha inizialmente sperimentato un momento di rilassamento passivo su un letto vibroacustico; questo specifico strumento terapeutico, convertendo la musica e le basse frequenze in onde meccaniche, trasmette vibrazioni mirate direttamente attraverso il corpo. Tale stimolazione profonda dei recettori tattili e propriocettivi ha lo scopo di favorire la distensione muscolare, ridurre l'ansia e facilitare la regolazione emotiva. Terminata questa fase di ricezione sensoriale, il Soggetto 1 è passato a una fase di interazione attiva e gioco sociale. I grafici generati dalla fase di *pre-processing* riflettono in modo speculare questi due momenti distinti, evidenziando la coerenza tra il dato biologico invisibile e il comportamento annotato sul protocollo.

L'attività elettrodermica: regolazione sensoriale e interazione sociale

Osservando il primo riquadro, dedicato alla componente tonica dell'EDA (la linea rossa), si nota un andamento molto peculiare nella prima fase della sessione (dalle 17:10 alle 17:30). Il segnale parte da valori alti (intorno ai $2.0 \mu\text{S}$) e si mantiene in una linea piatta e stabile. In questo arco temporale, il diario clinico riporta che il soggetto è sdraiato sul letto vibroacustico e ne sperimenta la pressione profonda. Da un punto di vista fisiologico, questa linea piatta e alta non indica noia o assenza di stimoli, ma individua uno stato di regolazione sensoriale; grazie alle vibrazioni, il sistema nervoso si rilassa perché percepisce una sensazione di sicurezza e di piacevole accompagnamento.

Il momento cruciale della sessione si verifica alle 17:33: il tracciato tonico mostra un crollo improvviso verso lo zero (gestito e recuperato grazie alla correzione di *offset* dell'algoritmo), immediatamente seguito da un'esplosione nel segnale grezzo e nella componente fasica (riquadro 2), che registra il picco più alto e denso di *driver* nervosi di tutta la seduta. L'analisi combinata di questa apparente anomalia e della successiva

reazione biologica trova spiegazione nel protocollo, infatti proprio alle 17:33 il paziente si alza dal letto per iniziare a giocare, passandosi oggetti luminosi con il Soggetto 2. Il crollo a zero del segnale è un chiaro artefatto da distacco temporaneo del sensore causato dal brusco cambio posturale (da sdraiato in piedi). Tuttavia, l'enorme picco fasico successivo, che decade poi molto lentamente, è segnale biologico puro.

Questa reazione nervosa rappresenta l'esempio di ciò che Welch [5] definisce *social context switch*. Infatti, il Soggetto 1 passa da uno stato non-sociale e passivo a uno stato interattivo. Di conseguenza, per gestire questa interazione, il sistema simpatico si attiva immediatamente, causando un repentino aumento dell'EDA che non deve essere interpretato dall'algoritmo come un segnale di ansia o pericolo, bensì come una sana reazione del soggetto che sta semplicemente attivando le proprie risorse per affrontare la novità sociale in modo adeguato.

Il disaccoppiamento tra il movimento e l'attività cardiaca: l'assenza di stress acuto

La conferma che l'attivazione delle 17:33 sia puramente positiva e ludica emerge dall'analisi congiunta del terzo riquadro (l'accelerometro) e del quarto riquadro (la frequenza cardiaca). Osservando quanto acquisito dall'accelerometro dalle 17:33 in poi, il grafico esplose in un tracciato ampio e rumoroso, coincidendo con l'inizio del gioco fisico.

Tuttavia, osservando il grafico della frequenza cardiaca (HR), si nota un fenomeno estremamente importante: il disaccoppiamento tra cuore e movimento. Nonostante la presenza di un'attività motoria intensa e prolungata, la frequenza cardiaca non mostra aumenti repentini o la tachicardia prolungata tipiche dello stress acuto. La linea dei battiti per minuto si mantiene stabile e su valori medio-bassi. In accordo con gli studi di Lydon et al. [27] e Groden [14], se il soggetto stesse vivendo una crisi di panico o di rabbia, il cuore avrebbe accelerato drasticamente per preparare i muscoli allo sforzo. In questo caso, il fatto che si registri un movimento intenso associato a una frequenza cardiaca bassa e stabile dimostra che l'attività motoria non costa alcuno sforzo emotivo al soggetto: il corpo si muove e gioca, ma il cuore rimane calmo. Questa dinamica potrebbe dunque portare ad abbassare la probabilità di agitazione o crisi nel soggetto.

HRV e dominanza parasimpatica

Osservando l'ultimo riquadro relativo alla variabilità della frequenza cardiaca, è possibile notare un valore RMSSD medio di ben 97 ms il quale indica una forte dominanza vagale (parasimpatica). Da questo parametro si apprende che il soggetto ha raggiunto uno stato di profondo benessere e che la terapia vibroacustica ha funzionato. Leggendo questo dato alla luce della letteratura scientifica (come il modello di Fenning e colleghi [15]), si ha la prova clinica del perché non è presente alcun rischio di crisi o di aggressività. Un valore RMSSD a 97 ms è un risultato altissimo, che ammortizza naturalmente lo stress. Anche alle 17:33, nel momento di maggiore attivazione fisica dovuta al gioco, il sistema nervoso parasimpatico, responsabile del rilassamento, ha mantenuto un'attività costante, gestendo l'intensità emotiva con grande naturalezza. La terapia ha generato nel paziente uno stato di stabilità e serenità così profondo da escludere il rischio di una crisi comportamentale.

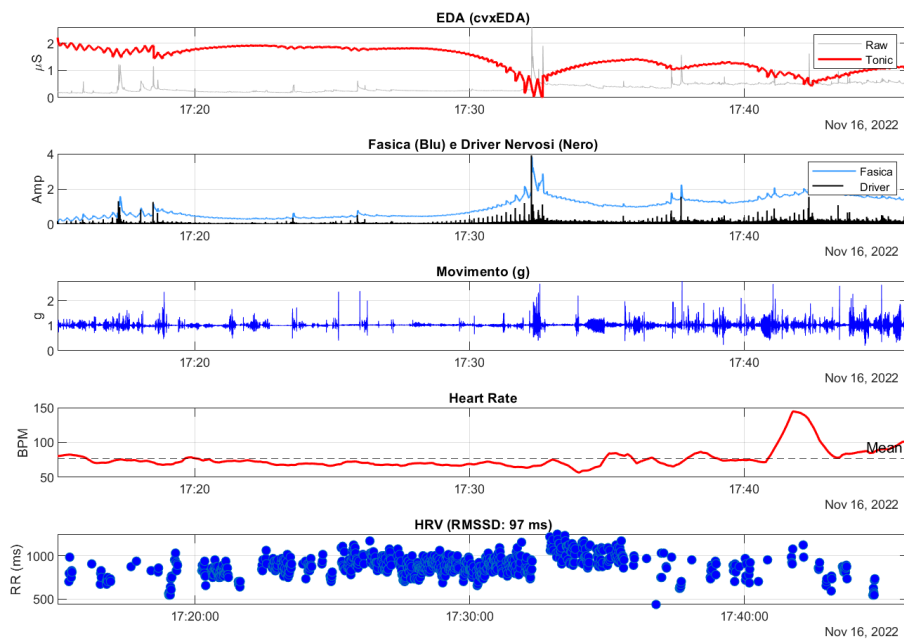


Figura 5.1: Dashboard multi-parametrica del Soggetto 1 durante la sessione 1 (16 novembre). I grafici documentano la transizione tra la regolazione sensoriale passiva e l'interazione attiva. Nella prima fase (fino alle 17:30), l'EDA tonica mostra un *plateau* elevato che indica rilassamento indotto dal letto vibroacustico. Alle 17:33 si osserva un marcato *social context switch*: all'inizio del gioco fisico, corrispondono un'esplosione di *driver* fasici (riquadro 2) e di movimento motorio (riquadro 3). Il disaccoppiamento con la frequenza cardiaca, che si mantiene bassa e stabile (riquadro 4), unito all'elevato valore di variabilità cardiaca (RMSSD: 97 ms, riquadro 5), conferma che questa attivazione è puramente ludica e totalmente priva di stress acuto.

Sessione 2: Analisi della sessione vibroacustica

La registrazione di questa seconda sessione ha presentato una notevole sfida tecnica, a causa delle vibrazioni meccaniche del letto che disturbano fisiologicamente i sensori. Nonostante ciò, l'analisi dei dati restituisce un risultato accurato a livello temporale, permettendo di documentare in modo chiaro l'effetto della stimolazione meccanica passiva sulla fisiologia del soggetto.

L'attività elettrodermica: il plateau omeostatico

Dalla componente tonica dell'EDA, è possibile notare un comportamento definibile come *plateau* omeostatico. In questa fase, infatti, il livello di conduttanza cutanea si mantiene quasi costante, fluttuando in modo minimo attorno al valore di $0.25 \mu S$ per l'intera durata della stimolazione principale (dalle 11:47 alle 11:58). Questa stabilità indica che il soggetto non sta subendo né un innalzamento dei livelli di stress né una disattivazione tipica della sonnolenza; la stimolazione vibroacustica induce piuttosto uno stato di sospensione fisiologica, mantenendo il sistema nervoso in un equilibrio vigile ma totalmente passivo. È inoltre rilevante notare come l'algoritmo di pre-elaborazione abbia filtrato con successo il rumore meccanico continuo, originariamente visibile come una seghettatura nel segnale grezzo, restituendo un tracciato tonico pulito.

L'analisi della componente fasica (riquadro 2) conferma questa stabilità di fondo: per gran parte della sessione la linea che traccia gli impulsi nervosi è piatta, dimostrando che le vibrazioni continue non vengono lette dal corpo come stimoli di stress, ma integrate come uno sfondo neutro. Si registrano unicamente due attivazioni fasiche distinte: un picco netto alle 11:50, in corrispondenza dello spostamento della vibrazione alla zona toracica, e un *cluster* di picchi tra le 11:53 e le 11:54, coincidente con una stimolazione estesa a tutto il corpo. L'evento delle 11:50 assume una particolare rilevanza clinica poiché coincide con l'intervento dell'operatore, che manipola l'attrezzatura in prossimità del corpo del soggetto, innescando una situazione che, di norma, genera tensione o disagio, poiché comporta l'ingresso di un'altra persona nello spazio intimo del soggetto [5].

Eppure, questa attivazione fisiologica si esaurisce nel giro di pochissimi istanti. Si

può dedurre che non ci si trova di fronte a una reazione di difesa o di allarme, ma a quella che in letteratura viene definita una sana "risposta di orientamento" [13]. Il Soggetto, in sostanza, si accorge della presenza dell'operatore e registra lo stimolo fisico, ma non lo avverte in alcun modo come una minaccia. Quanto osservato dimostra come l'ambiente protetto della stanza *Snoezelen* abbia trasformato profondamente il modo in cui il soggetto percepisce e difende il proprio spazio intimo. Un avvicinamento e un'interazione che, in una situazione ordinaria, avrebbero rischiato di innescare una crisi o una reazione oppositiva, in questo contesto vengono accolti con totale sicurezza e tolleranza.

Il disaccoppiamento tra il movimento e l'attività cardiaca: la stimolazione passiva

Il tracciato accelerometrico si presenta coerente con la posizione supina del soggetto, mostrando un'attività basale interrotta dai picchi delle 11:50 e 11:53, perfettamente sincroni con gli aggiustamenti posturali indotti dall'operatore.

Come già osservato nella prima sessione, questa dinamica ripropone il disaccoppiamento fisiologico-comportamentale [27, 14], ma in una veste nuova. In questo caso, i sensori inerziali registrano un input fisico importante dovuto al tremore del letto, che tuttavia non comporta il coinvolgimento del sistema cardiovascolare. Nonostante l'intensa energia meccanica esterna, la frequenza cardiaca rimane bassa e stabile, confermando l'assenza di un'attivazione da stress interno. Gli unici aumenti della frequenza cardiaca si osservano intorno alle 11:52 e alle 11:57, in risposta diretta alle stimolazioni fisiche più intense sul torace, e non configurano uno stato di *distress*. La rapidità con cui il battito ritorna ai valori basali, registrando un crollo a 70 bpm già alle 11:54, dimostra una elevata capacità di recupero vagale. Il cuore del soggetto risponde allo stimolo fisico acuto e recupera immediatamente, validando la natura puramente passiva e non stressogena della terapia.

L'HRV e la stabilità parasimpatica

L'analisi della variabilità della frequenza cardiaca presenta frammentazioni attese, causate dall'inevitabile interferenza della vibrazione meccanica con il sensore ottico al polso. Tuttavia, le porzioni di tracciato valide restituiscono un parametro RMSSD medio di 92 ms, un valore elevato e del tutto coerente con i 97 ms registrati nella prima

sessione, indicatore di una stabile dominanza parasimpatica. Questo parametro conferma pienamente quanto emerso nella prima sessione, ma aggiunge un elemento clinico fondamentale: la stabilità nel tempo. Dimostra, infatti, che il soggetto riesce a mantenere il proprio equilibrio anche di fronte a una stimolazione fisica particolarmente intensa. Si può dunque arrivare alla seguente deduzione: la terapia vibroacustica fa molto di più che rilassare i muscoli sul momento; abitua il sistema nervoso a mantenere la calma, costruendo giorno dopo giorno una difesa solida e duratura contro le crisi comportamentali.

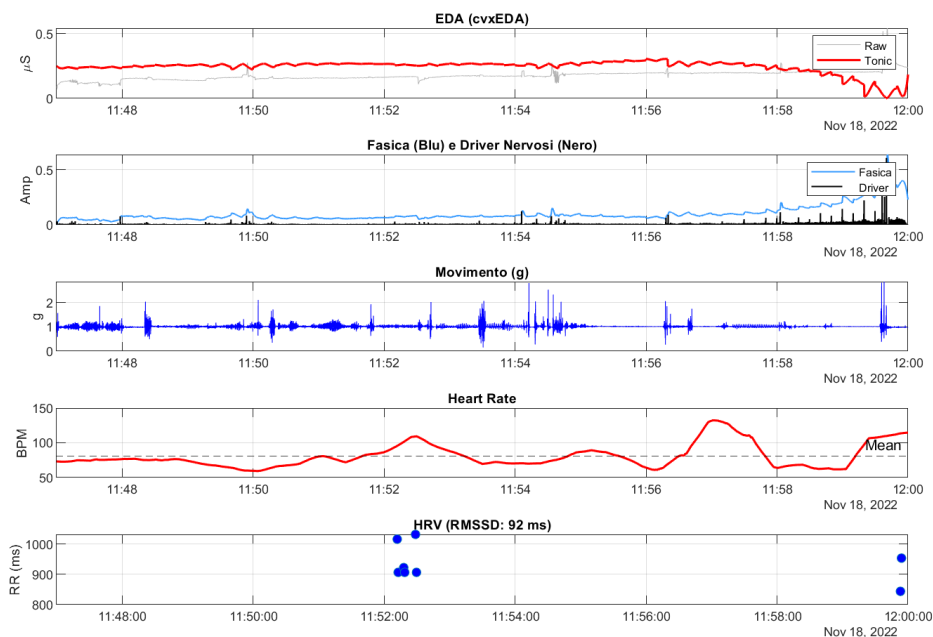


Figura 5.2: Dashboard multi-parametrica del Soggetto 1 durante la sessione 2 (18 novembre). I grafici documentano gli effetti della stimolazione vibroacustica passiva. Nel primo riquadro si osserva un *plateau* omeostatico dell'EDA tonica, indice di uno stato di sospensione fisiologica privo di allarme, ma comunque vigile. L'attività fasica è piatta, eccezion fatta per brevi e sane "risposte di orientamento" (es. 11:50) dovute all'intervento dell'operatore. Il disaccoppiamento tra l'intensa attività motoria (registrata dall'accelerometro a causa delle vibrazioni meccaniche, terzo riquadro) e la frequenza cardiaca, che si mantiene stabile (quarto riquadro), conferma l'assenza di *distress*. L'ultimo riquadro mostra la frammentazione del segnale HRV causata dalle inevitabili interferenze meccaniche con il sensore, ma restituisce un elevato parametro RMSSD (92 ms), confermando una stabile e prolungata dominanza parasimpatica.

5.6.2 Il caso del Soggetto 2

Mentre il profilo del Soggetto 1 ha permesso di osservare l'impatto del rilassamento profondo, il caso del Soggetto 2 esplora una dinamica clinica complementare: la gestione fisiologica dell'attenzione e dell'interazione attiva (*engagement*). L'esame delle due sessioni permetterà di valutare la reazione del sistema nervoso autonomo agli stimoli ambientali e visivi, dimostrando come i biosegnali consentano di distinguere in modo netto uno stato di genuino interesse esplorativo da un potenziale rischio di sovraccarico sensoriale.

Sessione 1: Baseline sensoriale

Durante questa sessione, il soggetto è stato inserito in un ambiente multisensoriale (stanza *Snoezelen*) e lasciato libero di interagire con oggetti luminosi. Si evidenzia come l'analisi visiva della *dashboard* multiparametrica generata dall'algoritmo restituisca un quadro clinico coerente con le dinamiche annotate sul diario clinico.

L'attività elettrodermica: l'engagement

Osservando il primo riquadro, relativo al confronto tra il segnale EDA grezzo e la sua componente tonica (la linea rossa), il primo fenomeno che salta all'occhio è di natura puramente tecnica. Intorno alle 16:20, momento in cui il braccialetto viene allacciato al polso, il tracciato mostra un crollo verticale seguito da una risalita molto ripida tra le 16:24 e le 16:28. Questo transitorio non rappresenta un'anomalia clinica, ma è il classico *settling time* (tempo di assestamento), ovvero il sensore necessita fisiologicamente di 5-10 minuti affinché la pelle crei il microclima umido sotto gli elettrodi necessario per stabilizzare la lettura.

Superata questa prima fase (dalle 16:30 in poi), la componente tonica si assesta su un *plateau* alto e costante. Questo livello elevato indica che il soggetto non è isolato o non reattivo, ma si trova in uno stato di forte *engagement* (coinvolgimento). Ciò significa che è mentalmente attivo e sta interagendo con gli oggetti luminosi. Dall'integrazione con lo stato dell'arte, questo grafico rappresenta l'esempio del "paradosso dell'oggetto" descritto da Prince et al. [12]. Se l'algoritmo valutasse questo livello alto di sudorazione decontestualizzato durante un compito cognitivo imposto, lo classificherebbe

rebbe erroneamente come rischio di crisi. Tuttavia, poiché il soggetto interagisce con un oggetto sensoriale e meccanico (le luci che cambiano colore), l'elevata conduttanza correla positivamente con comportamenti di interesse e auto-stimolazione (*stimming*), configurando uno stato di alto *arousal* a valenza positiva. Inoltre, in linea con gli studi di McCormick et al. [26], il fatto che la linea tonica mantenga un *plateau* e non mostri una pendenza in crescita infinita suggerisce che il sistema nervoso ha accettato lo stimolo e lo sta elaborando in modo sano, senza andare in sovraccarico (assenza di *failure to habituate*).

Il secondo riquadro della *dashboard* dimostra l'efficienza dell'algoritmo cvxEDA. Infatti, nonostante il segnale grezzo fosse inizialmente basso, il sistema ha estratto in modo nitido l'attività fasica e i singoli *driver* nervosi (le linee nere). L'allineamento con il protocollo è evidente: alle 16:25 si osserva un enorme picco (un *cluster* denso di impulsi neri) che coincide esattamente con il momento in cui il diario riporta che "il soggetto sperimenta con oggetti luminosi". Si tratta dunque della reazione fisiologica alla novità, cioè l'oggetto introdotto. Successivamente, tra le 16:53 e le 17:02, emergono picchi netti e isolati, corrispondenti al momento in cui il paziente riprende l'interazione libera con le fibre ottiche.

La struttura morfologica di questi picchi (alta frequenza ma ampiezza media) è classificata da Greenlee et al. [10] come il *pattern* tipico dell'autonomia e del coinvolgimento esplorativo. Se il soggetto fosse stato forzato a toccare le luci, si sarebbero osservati meno picchi ma di ampiezza decisamente maggiore (indice di *distress*). Inoltre, questa vivace reattività agli stimoli permette di escludere categoricamente l'ipotesi di ansia cronica "smussata" (*blunted reactivity*) o di *shutdown* descritta da Panju et al. [11]. Il sistema simpatico del Soggetto 2 risponde dunque in modo appropriato e dinamico.

L'attività motoria: esplorazione tattile e ricerca di stimolazione

Il terzo riquadro, dedicato all'accelerometro (ACC), mostra un tracciato in cui, a differenza di una motricità grossolana che genererebbe onde continue, si osservano lunghe fasi di segnale piatto interrotte da picchi verticali rapidi e distinti. Il segnale strutturato in tal modo, permette all'algoritmo di identificare con precisione la motricità fine,

ovvero la manipolazione mirata degli oggetti con le mani. L'aspetto più rilevante è la sincronia tra questi picchi di movimento e l'innalzamento dell'EDA, dimostrando il legame puro tra attivazione elettrodermica e azione. Inoltre, come teorizzato da Lydon et al. [27], il Soggetto 2 non si sta muovendo per scaricare la tensione e calmarsi, ma utilizza il movimento ripetitivo per ricercare e mantenere attivamente una stimolazione sensoriale gratificante. Tuttavia, dal punto di vista predittivo, questa è una condizione importante a cui prestare attenzione; infatti, la Teoria dell'Interferenza [4] insegna che l'interruzione brusca di questi movimenti ripetitivi da parte di un operatore rappresenta una potente causa per l'insorgere di un comportamento aggressivo. L'algoritmo deve tenere conto quindi di questa condizione, etichettandola come stabile ma potenzialmente vulnerabile in caso di interruzione dall'esterno.

L'attività cardiaca: equilibrio e stress positivo

Gli ultimi due riquadri chiudono l'analisi confermando il benessere generale del soggetto attraverso l'osservazione dell'attività cardiaca. Nel quarto riquadro, il tracciato della frequenza cardiaca (HR) mostra valori estremi nei primissimi minuti (16:20-16:25), riconducibili allo stesso artefatto di contatto visto nell'EDA: il sensore ottico si stava ricalibrando durante l'allacciamento del cinturino. Raggiunta la stabilità, il battito si assesta su una media di 85-90 BPM. Questo valore è perfettamente fisiologico per uno stato di veglia attiva, escludendo sia condizioni di tachicardia da stress acuto (che supererebbero i 110 BPM) sia bradicardie. Questo vuol dire che il sistema cardiaco sta semplicemente sostenendo uno stato di stress positivo, in altre parole, funzionale al mantenimento dell'attenzione.

Osservando il grafico relativo all'HRV, il valore calcolato per la metrica RMSSD è di 53 ms. Una simile densità dei punti e un valore così elevato indicano una capacità di regolazione vagale molto buona. Dunque, non vi è traccia di stress acuto (che farebbe crollare l'RMSSD sotto i 20 ms), ma si evidenzia uno stato di calma vigile. Riconducendo questo caso al modello teorico di Fenning et al. [15], ci si trova di fronte al classico "profilo misto protettivo". Nonostante l'alta attivazione del sistema simpatico (EDA elevata legata all'interesse), il paziente mantiene un solido "freno vagale" (RMSSD a 53 ms). In termini pratici, l'organismo è in uno stato di eccitazione dovuto allo

stimolo ma si sente al sicuro e perfettamente autoregolato.

Alla luce di questi dati, si può affermare che il pericolo di una crisi comportamentale in questa fase è pressoché inesistente giustificando una semplice attivazione positiva legata ad un momento sano di gioco da parte dell'individuo.

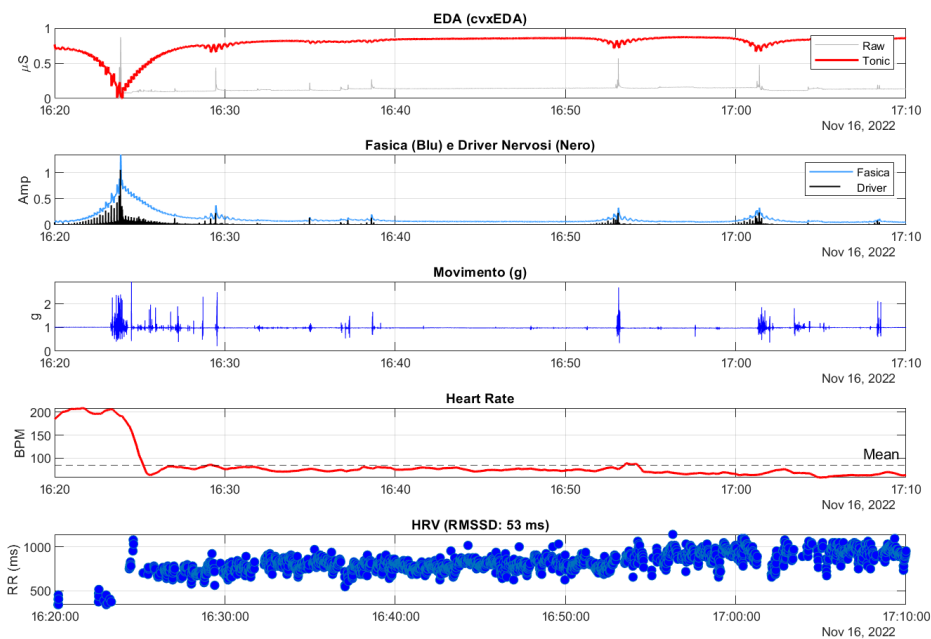


Figura 5.3: Dashboard multi-parametrica del Soggetto 2 durante la Sessione 1 (baseline sensoriale). Nel primo riquadro, superato il transitorio iniziale di *Settling Time* (fino alle 16:28), l'EDA tonica mostra un *plateau* elevato che indica un forte stato di *engagement* sensoriale, non di ansia. Il secondo e terzo riquadro evidenziano una netta sincronia tra i picchi fascici nervosi e i micromovimenti, corrispondenti ai momenti di interazione diretta e autonoma con gli oggetti luminosi (es. 16:25 e 16:53-17:02). La frequenza cardiaca stabile (riquadro 4) e il buon valore di regolazione vagale (RMSSD di 53 ms, riquadro 5) confermano uno stato di "stress positivo" e calma vigile, escludendo del tutto il rischio di *distress*.

Sessione 2: Sessione sensoriale con luce di Wood

La seconda sessione analizzata, registrata il 18 novembre, presenta il caso in cui il soggetto è stato esposto a un ambiente visivamente caratterizzato dall'uso della luce di Wood e di materiali fluorescenti. Se la prima sessione ha permesso di rilevare tramite i segnali processati un'alta attivazione, questa seconda acquisizione prevede la decodifica di segnali biologici più "deboli" e silenziosi, restituendo l'immagine di una mente profondamente concentrata.

Prima di addentrarsi nell'analisi dei segnali fisiologici, è doveroso chiarire la natura dello stimolo ambientale a cui il soggetto è stato esposto, dal momento che, come già ampiamente affrontato, il contesto in cui il soggetto si trova è strettamente influente al suo comportamento e coinvolgimento emotivo. La lampada di Wood, nota comunemente come "luce nera", è una sorgente luminosa che emette radiazioni elettromagnetiche quasi esclusivamente nella gamma degli ultravioletti a onda lunga (UV-A), producendo una quantità di luce visibile per l'occhio umano estremamente ridotta. La caratteristica fondamentale di questa luce è la sua capacità di indurre il fenomeno della fluorescenza: quando i raggi UV invisibili colpiscono determinati materiali appositamente preparati per le terapie (come plastiche o tessuti), questi assorbono l'energia e la riemettono istantaneamente sotto forma di luce visibile, accendendosi di colori vividi e brillanti.

Nel contesto del disturbo dello spettro autistico (ASD) e delle stanze multisensoriali, questo strumento persegue due obiettivi clinici precisi. Il primo è legato alla generazione di un totale isolamento visivo dovuto al fatto che, affinché l'effetto fluorescente sia visibile, la stanza deve essere immersa nell'oscurità. Dunque, questa condizione di buio elimina in un solo istante tutto ciò che disturba il quadro visivo e distrae dall'ambiente circostante, come ad esempio i contorni della stanza, i mobili, i riflessi, che spesso causano sovraccarico sensoriale nel soggetto. In secondo luogo, questo strumento funge da potente catalizzatore per l'attenzione (*engagement*): in questo spazio oscurato, gli unici elementi visibili diventano gli oggetti fluorescenti. Questo contrasto cromatico estremo cattura lo sguardo e l'attenzione del soggetto, incoraggiandolo all'inseguimento oculare e alla manipolazione fine, stimolando la curiosità senza generare l'ansia tipica degli ambienti iper-stimolanti. Una volta chiarita la peculiarità del contesto ambientale, ci si può addentrare nella spiegazione e nello studio dei dati fisiologici registrati in questa sessione.

L'attività elettrodermica: il profilo low responder

Osservando la parte tonica e fasica dell'EDA, si nota immediatamente una differenza sostanziale rispetto alla *baseline* precedente. La scala dei valori assoluti è bassissima, intorno agli $0.1 \mu\text{S}$. Questo dato inquadra chiaramente il Soggetto 2 come

un fenotipo fisiologico *low responder*, ovvero un individuo la cui pelle oppone un'alta resistenza o conduce naturalmente poco, un fattore spesso legato a caratteristiche dermatologiche, come la secchezza cutanea, o a un'impedenza fisiologica dell'elettrodo. Tuttavia, l'algoritmo cvxEDA, rispetto ai classici filtri, grazie alla matematica del problema inverso, è riuscito ad estrarre dal microscopico segnale grezzo quattro eventi fasici distinti (evidenziati dai *driver* neri): uno all'inizio dell'esplorazione (11:21), uno molto pronunciato al minuto 11:34 e due verso la fine della sessione (11:38 e 11:43).

Dal punto di vista clinico, è fondamentale non confondere questo tracciato con il profilo di "reattività smussata" [11]. Nel caso dello *shutdown* ansioso documentato in letteratura, il segnale è piatto perché il sistema nervoso è collassato e non reagisce più all'ambiente. Al contrario, nel caso del Soggetto 2, la presenza di questi picchi nervosi netti e tempestivi dimostra che la competenza reattiva biologica è perfettamente intatta: il sistema autonomo "sente" lo stimolo fluorescente e risponde, seppur con un'ampiezza ridotta dalla resistenza cutanea.

L'attività motoria: sincronia con l'attività cardiaca ed engagement sensoriale

La prova definitiva che i picchi di sudorazione estratti non siano semplici artefatti numerici o rumore di fondo arriva dall'incrocio con il terzo riquadro (l'accelerometro). L'asse del movimento si presenta quasi come una linea piatta (indicando che il paziente è seduto e tranquillo), interrotta unicamente da brevi istanti di motricità fine. Questi micromovimenti di manipolazione degli oggetti fluorescenti coincidono esattamente al millisecondo con i quattro picchi fasici dell'EDA.

Ancora più rivelatore è il comportamento del cuore, tracciato nel quarto riquadro. Dopo la fisiologica fase di assestamento dei primi due minuti (un picco sopra i 100 BPM dovuto all'ingresso nella stanza e al posizionamento), la frequenza cardiaca scende e raggiunge il suo naturale ritmo di base, stabilizzandosi in una fascia estremamente regolare tra i 75 e gli 85 BPM. In questa fase, il sistema cardiovascolare entra in una condizione di "riposo attivo", in altre parole, il soggetto è sveglio, vigile e partecipa, ma il suo organismo non è sottoposto ad alcuno sforzo metabolico o tensivo.

Quello che si evince è che esattamente nei momenti in cui il soggetto muove le mani per manipolare gli oggetti e la sudorazione segna un picco (11:34 e 11:38), a li-

vello cardiaco rimane perfettamente calmo e non subisce alcuna accelerazione. Questo dettaglio è cruciale per la corretta interpretazione dei dati: in una normale reazione di allarme, di paura o di stress acuto, il sistema nervoso simpatico attiverebbe l'intero organismo facendo inevitabilmente accelerare anche il battito cardiaco per preparare il corpo a una reazione di difesa. In questo caso, invece, l'imperturbabilità del cuore dimostra che il picco di sudorazione non è assolutamente derivato dall'ansia o dal nervosismo, ma rappresenta la pura, serena e focalizzata attenzione mentale verso l'oggetto che si sta esplorando.

Dunque, questa combinazione specifica tra micromovimenti e picchi fasici in totale assenza di tachicardia, classifica l'evento, secondo il modello di Prince et al. [12], come puro *engagement* sensoriale. Grazie alla presenza dello stimolo fornito dall'oggetto fluorescente, il movimento è finalizzato a esplorare fisicamente il materiale e il picco di conduttanza riflette la semplice attivazione emotiva e positiva della scoperta. Inoltre, l'assoluta mancanza di quel lento accumulo e innalzamento tonico descritto da Ferguson et al. [9] (*anticipatory rise*) conferma che l'attivazione è episodica e positiva, e non rappresenta in alcun modo un accumulo di tensione pre-aggressiva.

L'HRV come indicatore di concentrazione

Il valore calcolato di RMSSD per questa sessione è di 43 ms. Rispetto ai 53 ms registrati nella prima *baseline*, si nota un lieve abbassamento, ma il valore si mantiene all'interno di un *range* fisiologico sano. Per capire il reale significato clinico di questo lieve calo a 43 ms, ci si può affidare agli studi di Fenning et al. [15] e alla teoria polivagale. In termini molto pratici, questo valore numerico rappresenta uno stato di "attenzione sostenuta". Ovvero, quando una persona è fortemente concentrata a osservare qualcosa di specifico, come in questo caso i colori accesi generati dalla luce di Wood, il suo corpo non può rimanere in uno stato di abbandono o rilassamento totale, ma, per permettere alla mente di focalizzarsi visivamente sugli oggetti, il sistema nervoso deve fisiologicamente ridurre un po' il suo livello di quiete profonda (un meccanismo di adattamento noto in clinica come *vagal withdrawal*, ovvero il leggero ritiro del freno vagale).

La differenza diventa ancora più chiara e immediata se si effettua un paragone con

i dati di un altro paziente osservato in questo stesso studio: durante una sessione di riposo assoluto e passivo, il Soggetto 1 ha registrato un valore RMSSD di ben 97 ms. Il dato di 43 ms del Soggetto 2 descrive quindi una situazione diversa, ma ugualmente sana e funzionale. Non ci si trova di fronte a uno stato di stress, di ansia o di pericolo, ma semplicemente allo sforzo naturale e necessario che il cervello compie per elaborare le informazioni.

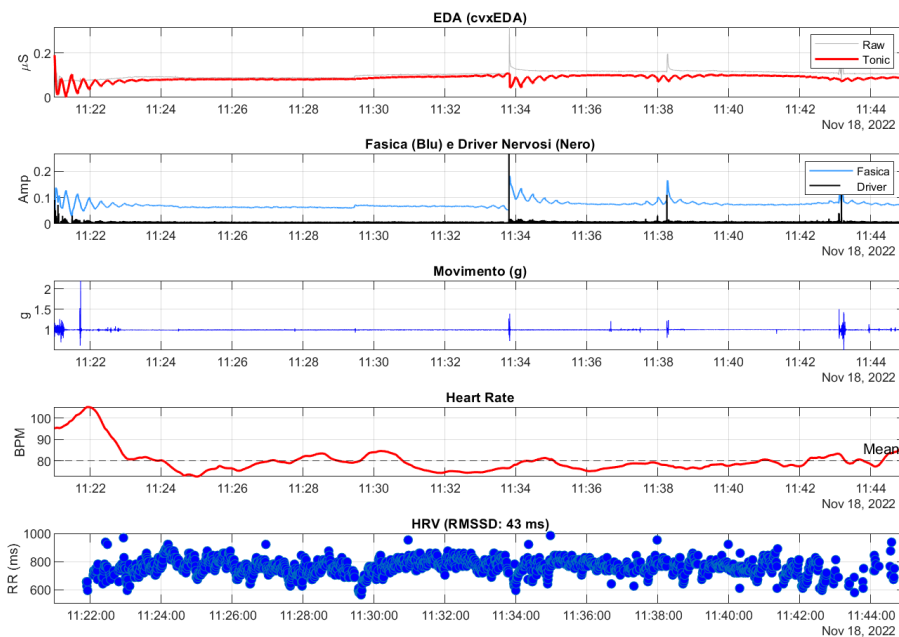


Figura 5.4: Dashboard multi-parametrica del Soggetto 2 durante la Sessione 2 (luce di Wood). Il primo riquadro evidenzia un fenotipo *low responder* (valori EDA assoluti molto bassi, $\sim 0.1 \mu S$). Tuttavia, l'algoritmo estrae chiari picchi fasici (es. 11:34 e 11:38) che risultano perfettamente sincroni con i micromovimenti esplorativi delle mani (riquadro 3). La totale assenza di tachicardia in corrispondenza di tali attivazioni (riquadro 4, HR stabile sui 75-85 BPM) esclude reazioni di allarme o stress, configurando un quadro di puro *engagement* sensoriale. Il valore RMSSD di 43 ms (riquadro 5) riflette uno stato di "attenzione sostenuta" e il fisiologico ritiro vagale necessario per focalizzarsi sugli stimoli fluorescenti.

5.6.3 Il caso del Soggetto 3

Il profilo del Soggetto 3 offre una prospettiva clinica diversa e particolarmente significativa rispetto ai casi precedenti. Questa sezione mette a confronto due distinte modalità di riduzione dell'attività fisiologica: una chimica e una fisica. La prima sessione documenterà la complessa fase di recupero del corpo da un acuto attacco di panico a

seguito di una sedazione farmacologica; la seconda analizzerà l'impatto di una tecnica fisica mirata, la pressione profonda, capace di indurre uno stato di calma altrettanto forte ma mantenendo il paziente perfettamente vigile e connesso all'ambiente. Questo confronto diretto permetterà di oggettivare la profonda differenza biologica che intercorre tra la sedazione indotta e una reale e sana regolazione emotiva.

Sessione 1: Analisi della fase di recovery post-panico

L'analisi visiva della *dashboard* multiparametrica permette di suddividere la sessione in tre fasi ben distinte, che riflettono in modo fedele e sequenziale l'effetto farmacodinamico della sedazione e la progressiva disattivazione del sistema nervoso simpatico.

L'attività elettrodermica: lo stress residuo e l'effetto del farmaco Nei primi minuti di registrazione (dalle 17:48 alle 17:55) si nota immediatamente un picco di proporzioni elevate, visibile sia nella componente tonica (la linea rossa) sia nella fitta presenza di *driver* nervosi fasici (gli impulsi neri). Fisiologicamente, il corpo impiega circa 20-30 minuti per smaltire i livelli ematici di cortisolo e adrenalina rilasciati durante una crisi. Dunque, quello che è visibile nel grafico è proprio la "coda" dell'attacco di panico appena vissuto, sommata allo stress residuo legato all'azione di indossare il braccialetto. Nonostante il Soggetto sia ormai in una situazione di sicurezza, il suo sistema nervoso autonomo è ancora inondato dai residui ormonali della crisi.

Secondo Ferguson et al. [9], l'incapacità del corpo di tornare rapidamente ai valori di base dopo un evento stressante è uno dei più forti predittori di una possibile recidiva o di un nuovo scoppio aggressivo, a causa dell'accumulo di tensione. In questo caso, però, l'intervento farmacologico modifica drasticamente questa naturale e alquanto lenta curva di recupero.

A partire dalle 17:55, infatti, il tracciato cambia la sua morfologia. Fino alle 18:19, la linea che registra gli impulsi nervosi si appiattisce a causa dell'effetto diretto della benzodiazepina (Tavor). Il farmaco agisce sui recettori GABAergici e inibisce il sistema simpatico prima ancora che il cortisolo venga smaltito in modo naturale: Il soggetto è cosciente, ma il farmaco ha disattivato la sua risposta fisiologica allo stress." Si noti

come questa fase di sedazione farmacologica assomigli visivamente al *pattern* di reattività smussata [11] nei soggetti affetti da ansia cronica severa. Tuttavia, in questo caso l'assenza di segnale nervoso non è dovuta a un esaurimento funzionale o a un collasso emotivo, ma unicamente a un blocco chimico indotto dall'esterno.

L'attività motoria e cardiaca: immobilità, sedazione e riaccoppiamento

L'azione sedativa del farmaco trova piena conferma nei dati registrati dall'accelerometro e dal sensore cardiaco. Infatti, tra le 17:50 e le 18:19, si osserva che il tracciato del movimento è una linea sottilissima, pressoché piatta. Si può dedurre che il soggetto è seduto o sdraiato, immobile, probabilmente esausto dopo la crisi di panico e profondamente rilassato dall'ansiolitico. Questa immobilità si riflette anche sul comportamento cardiaco: anziché registrare una tachicardia fissa, tipica di chi sta ancora subendo lo stress, il battito cardiaco oscilla in modo ampio e regolare tra i 60 e i 100 BPM. Questo dimostra che il farmaco ha permesso al sistema parasimpatico di riprendere il controllo, forzando un ingente rilassamento cardiaco e muscolare e riducendo in modo significativo il carico di lavoro del cuore.

La situazione muta di nuovo intorno alle 18:19. Sul grafico riappaiono i picchi nervosi dell'EDA, l'accelerometro registra un'esplosione di movimento, probabile segno che il paziente si sta alzando per uscire dalla stanza, e la frequenza cardiaca aumenta, toccando quasi i 140-150 BPM dimostrando che il sistema cardiovascolare è rimasto reattivo. Non appena la richiesta di energia aumenta (perché il paziente si alza e cammina), il sistema cardiaco risponde immediatamente allo sforzo fisico, uscendo dallo stato di sedazione profonda. Questa risposta tempestiva certifica un riaccoppiamento fisiologico sano [27], confermando che il sistema nervoso del soggetto era unicamente sedato dal farmaco, ma non presentava alcuna disfunzione strutturale.

L'HRV: il recupero del sistema parasimpatico e il ritorno all'equilibrio

Il valore RMSSD calcolato per questa sessione è di 91 ms, un valore altissimo che, fisiologicamente, descrive la risposta compensatoria del sistema nervoso. Dopo un attacco di panico, che è un evento dominato dal sistema simpatico e dall'adrenalina, il corpo interviene mediante il sistema parasimpatico per calmare il cuore e ripristinare l'equilibrio. Il farmaco ansiolitico somministrato non fa altro che favorire e accelerare

questo naturale processo di rilassamento.

Un valore RMSSD a 91 ms, unito al crollo totale dell'attività elettrodermica visto in precedenza, posiziona il paziente perfettamente all'interno del profilo di "basso rischio" descritto da Fenning et al. [15]. Questo conferma il successo della terapia con la transizione dal pericoloso stato di allarme (alta EDA e basso HRV, tipico del panico) a uno stato altamente protettivo (bassa EDA e alto HRV), ripristinando in pieno l'efficacia del freno vagale.

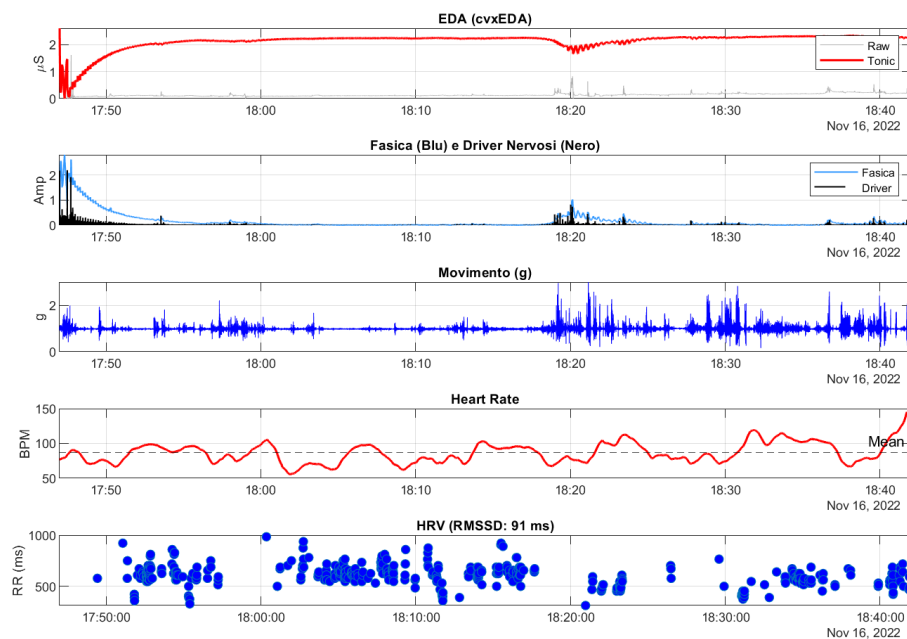


Figura 5.5: Dashboard multi-parametrica del Soggetto 3 durante la Sessione 1 (*recovery* post-panico). Il tracciato mostra tre fasi distinte. Inizialmente (fino alle 17:55), l'elevata attività elettrodermica tonica e fasica riflette lo stress ormonale residuo della crisi. Nella fase centrale (17:55-18:19), è evidente l'effetto del farmaco sedativo: totale appiattimento dell'EDA (blocco simpatico) e immobilità motoria. Parallelamente, il cuore oscilla regolarmente e l'HRV registra un altissimo valore RMSSD (91 ms), confermando il recupero parasimpatico forzato e il raggiungimento di un "profilo di basso rischio". Infine, dalle 18:19, la brusca ripresa di movimento, sudorazione e frequenza cardiaca documenta l'alzata del paziente, certificando un sano riaccoppiamento fisiologico.

Sessione 2: Analisi della sessione con pressione profonda

I grafici di questa seconda sessione documentano l'effetto della stimolazione somatica tramite pressione profonda. A differenza della prima seduta, dove la calma era indotta

dalla somministrazione di un farmaco, i dati che emergono in questa fase descrivono la naturale capacità del corpo di autoregolarsi grazie all'input fisico esterno.

I primissimi minuti della sessione (dalle 10:35 alle 10:45) rappresentano la fase di preparazione; motivo per cui il tracciato dell'attività elettrodermica mostra un denso gruppo iniziale di picchi nervosi fasici, perfettamente coerenti con i movimenti registrati dall'accelerometro. Si tratta della fisiologica e sana attivazione legata all'arrivo nella stanza, al mettersi in posizione e all'inizio dell'interazione con l'operatrice. Intorno alle 10:43 si registra un piccolo artefatto, rilevabile come un "buco" nel tracciato tonico, dovuto probabilmente a un movimento brusco o a un aggiustamento del sensore, che tuttavia l'algoritmo di pre-elaborazione ha riconosciuto e compensato.

L'attività elettrodermica: l'attenzione e l'accettazione del contatto

Nella fase centrale della seduta (dalle 10:50 alle 11:15), coincidente con il momento in cui l'operatrice applica la stimolazione somatica e la pressione profonda sul corpo del Soggetto 3, si osserva che i picchi fasici non sono presenti; tuttavia, a differenza di quanto accaduto durante la sedazione farmacologica della prima sessione (dove il segnale crollava verso lo zero), qui la componente tonica di base sale dolcemente e si mantiene su valori alti e costanti. Questo andamento del segnale documenta lo stato di regolazione attiva innescata dalla pressione profonda.

Di norma, come evidenziato in letteratura, un innalzamento continuo del livello tonico preannuncia un sovraccarico di tensione o uno scoppio aggressivo; in questo caso, però, l'aumento avviene in un contesto in cui i picchi fasici sono silenti: l'assenza di scatti improvvisi dimostra che il soggetto non sta reagendo a una minaccia e l'alta conduttanza rappresenta un sano sostegno metabolico per mantenere viva l'attenzione. Questa è la prova biologica dell'accettazione dello stimolo.

Questa deduzione assume un valore ancora maggiore se letto attraverso gli studi di Welch [5] e Verneti et al. [13]. Nei soggetti con ASD, il contatto fisico da parte di un operatore rappresenta spesso un'invasione critica dello spazio personale, innescando reazioni di allarme e difesa. Qui, l'assoluta mancanza di picchi fasici dimostra che la tecnica della pressione profonda è riuscita a bypassare questo sistema di allarme così che il tocco dell'operatrice non viene processato come una minaccia ma come un input

propriocettivo sicuro.

L'attività motoria e cardiaca: sincronia tra corpo e battito

L'efficacia di questa terapia trova piena conferma incrociando i dati del movimento con quelli della frequenza cardiaca (HR). Durante i 25 minuti di stimolazione profonda, il tracciato dell'accelerometro è una linea piatta, a conferma del fatto che il Soggetto si trova in uno stato di immobilità, sedato in modo naturale e pronto a ricevere la pressione.

Parallelamente, la frequenza cardiaca si stabilizza su valori di assoluto riposo, escludendo sia tachicardie (ansia repressa) sia bradicardie (addormentamento). A differenza dei profili di *distress* descritti da Lydon et al. [27], dove un corpo apparentemente fermo nasconde un battito accelerato per un'ansia inespressa, qui si osserva un perfetto sincronismo tra immobilità fisica e battito regolare. Il soggetto non è bloccato dal disagio (*freezing*), ma sta vivendo un profondo e reale rilassamento muscolare indotto dal corpo stesso. A fine sessione (intorno alle 11:20), non appena il braccialetto viene rimosso, i movimenti e i parametri si riattivano fisiologicamente, confermando la piena reattività del soggetto.

L'HRV: la conferma del benessere e l'assenza di allarme

A sigillare questo quadro di benessere interviene il dato della variabilità cardiaca (HRV), che registra un valore RMSSD di 74 ms.

Incrociando questa elevata attività parasimpatica, indice di calma e regolazione, con l'elevata attività tonica vista in precedenza, ovvero il sistema simpatico che sostiene l'attenzione, il soggetto può essere posizionato all'interno del "profilo misto protettivo" descritto da Fenning et al. [15]. Si tratta di co-attivazione preziosa grazie alla quale il soggetto è vigile e presente, ma allo stesso tempo si trova in uno stato di sicurezza e serenità.

In conclusione, l'analisi di questa sessione conferma l'efficacia della pressione profonda nell'inibire lo stress motorio e cardiaco, ma preservando la connessione del soggetto con l'ambiente, facilitando una sana regolazione emotiva senza ricorrere alla sedazione.

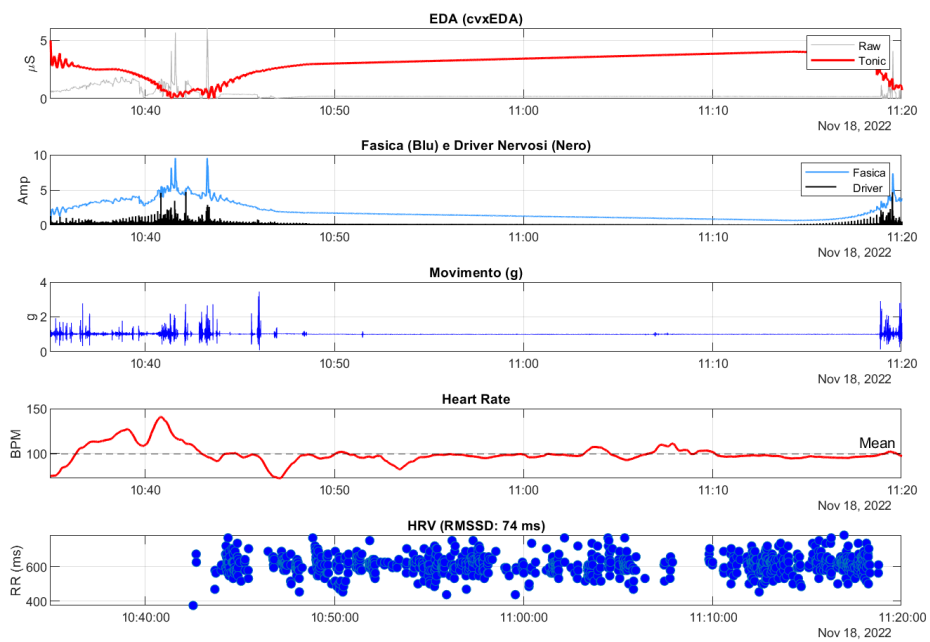


Figura 5.6: Dashboard multi-parametrica del Soggetto 3 durante la Sessione 2 (pressione profonda). Dopo una fisiologica attivazione iniziale (fino alle 10:45), la fase centrale di stimolazione (10:50-11:15) è caratterizzata dall'assoluta assenza di picchi fascici nervosi e da un dolce innalzamento dell'EDA tonica (riquadri 1 e 2). Questo *pattern* documenta la totale accettazione del contatto fisico e uno stato di sana regolazione attiva. Il perfetto sincronismo tra l'immobilità motoria (riquadro 3) e la stabilità della Frequenza Cardiaca su valori di riposo (riquadro 4) esclude stati di ansia repressa, evidenziando un profondo rilassamento muscolare naturale. Il valore RMSSD di 74 ms (riquadro 5) certifica il raggiungimento di un "profilo misto protettivo", in cui il paziente è contemporaneamente vigile, presente e serenamente connesso all'ambiente.

6. Estrazione delle feature

Dopo aver completato la fase di *pre-processing*, i segnali fisiologici risultano puliti da artefatti e scomposti nelle loro componenti fondamentali. Tuttavia, un segnale continuo nel tempo non può essere fornito direttamente a un algoritmo di *Machine Learning*, per questo motivo è necessario effettuare l'estrazione delle *feature*; procedura che consiste nel trasformare l'andamento continuo del segnale in un insieme di indicatori numerici discreti, le *feature* appunto, capaci di riassumere e quantificare specifici comportamenti fisiologici e clinici.

L'obiettivo di questo capitolo è illustrare come si è proseguiti con l'estrazione di questi biomarcatori. Il codice implementato non si basa su metriche scelte arbitrariamente, ma include i modelli presenti in letteratura per lo studio del sistema nervoso autonomo nel disturbo dello spettro autistico. Di seguito vengono analizzati, passo dopo passo, la logica matematica e clinica che guidano l'intero processo di estrazione.

6.1 Adattamento e calibrazione individuale

L'intero processo di estrazione prende avvio dall'acquisizione dei dati in ingresso. Nello specifico, l'algoritmo importa le matrici di dati (.mat) generati e salvati al termine della precedente fase di preelaborazione. Ogni file rappresenta un singolo segmento processato e contiene al suo interno i segnali fisiologici del soggetto (conduttanza cutanea, frequenza cardiaca e motricità) già filtrati, sincronizzati e privati degli artefatti principali.

Il primo e fondamentale requisito per garantire la riproducibilità scientifica dell'analisi è la corretta inizializzazione dell'ambiente di lavoro. L'algoritmo prevede l'isolamento sistematico del *workspace* a ogni avvio, assicurando che i dati estratti per un soggetto non vengano contaminati da residui di variabili appartenenti ad acquisizioni precedenti. Fatto ciò, il sistema carica i segmenti .mat e procede alla definizione dei parametri di base.

6.1.1 Definizione della *baseline*

Nell'analisi dei segnali fisiologici, la definizione dello "zero fisiologico", ovvero la *baseline*, è uno dei passaggi più decisivi. Questo parametro rappresenta lo stato di normale riposo del soggetto e funge da punto di riferimento poiché ogni successiva variazione del battito cardiaco o dell'attività elettrodermica assumerà un significato clinico (di allarme o di rilassamento) solo ed esclusivamente in relazione a questo valore di partenza.

Generalmente, la maggior parte degli studi presenti in letteratura utilizza un approccio standardizzato, assumendo in automatico i primi cinque o dieci minuti di registrazione come *baseline* di riposo per tutti i partecipanti. Tuttavia, questo approccio risulterebbe rigido e vincolante nell'ambito del disturbo dello spettro autistico rischiando di generare errori interpretativi rilevanti. Spesso, infatti, i soggetti giungono nella stanza di terapia già in uno stato di iper-attivazione (definito *arousal*) dovuto all'ansia per il nuovo ambiente, allo stress del viaggio o a un sovraccarico sensoriale precedente. Se l'algoritmo acquisisse un momento di forte stress come condizione di normalità, tutti i successivi calcoli risulterebbero sfalsati, mascherando eventuali crisi future. Per questo motivo, si è deciso di implementare una regola che si avvale dell'osservazione clinica del singolo soggetto e dell'analisi effettuata nel capitolo precedente, riconoscendo la specificità di ciascun soggetto e riducendo il rischio di falsi positivi e falsi negativi nel modello finale. In termini pratici, per ogni soggetto, è stato individuato l'esatto blocco temporale in cui l'individuo ha dimostrato di trovarsi nel suo reale stato di quiete, in particolare:

- **Il caso di Soggetto 1 - *baseline* di calma fisiologica:** È stata esclusa la fase iniziale (caratterizzata da una forte interazione sociale) in favore dell'intera sessione vibroacustica passiva. Come precedentemente validato, in questa finestra il soggetto ha raggiunto la massima dominanza vagale (RMSSD: 92 ms). Questo parametro certifica uno stato di profondo rilassamento e benessere. Calibrando l'algoritmo su questo stato di massima sicurezza, ogni successiva variazione o picco di stress risulterà immediatamente visibile e quantificabile.

- **Il caso di Soggetto 2 - *baseline* di quiete iniziale:** In questo caso si è adottato il metodo tradizionale dal momento che il protocollo e i tracciati hanno confermato che il soggetto ha mostrato un arrivo tranquillo, privo di alterazioni comportamentali o fisiologiche. Quindi è stato ritenuto appropriato mantenere come riferimento l'inizio standard della prima sessione.
- **Il caso di Soggetto 3 - *baseline* post-farmaco:** Come documentato in fase di preelaborazione, il soggetto, nella prima sessione, è arrivato trovandosi già all'interno di un attacco di panico pregresso, mostrando un'estrema attivazione del sistema nervoso simpatico. Chiaramente il sistema non avrebbe dovuto prendere quei primi minuti come riferimento di "normalità". A questo punto, ci si potrebbe chiedere perché non sia stata utilizzata come riferimento l'altra acquisizione disponibile, ovvero la sessione di lavoro con pressione profonda, durante la quale il soggetto non mostrava segni di stress ed era collaborativo. La motivazione risiede nel concetto stesso di "zero fisiologico". Infatti, durante la sessione di lavoro, il soggetto si trovava in uno stato di calma vigile, cioè il suo freno vagale era attivo, garantendo uno stato di calma, ma anche il suo sistema simpatico era attivo e l'EDA tonica era elevata per sostenere lo sforzo cognitivo e l'attenzione. Se l'algoritmo avesse acquisito questo stato di attivazione fisiologica come *baseline*, avrebbe impostato un valore di riferimento iniziale troppo alto, di conseguenza, sarebbe stato considerato normale uno stato di allerta, finendo per non riconoscere o sottostimare i futuri e reali scatti d'ansia. Dunque, si è scelto di estrapolare come valori di riferimento quelli della fase di recupero successiva alla somministrazione del farmaco sedativo (Tavor), momento esatto in cui il sistema nervoso del soggetto è stato chimicamente inibito, registrando un'attività simpatica silente. Questa assoluta inattività innescata fornisce l'unico e vero "zero assoluto" per questo soggetto, il punto di partenza più appropriato per misurare con precisione qualsiasi futura variazione.

6.1.2 Soglie di validità

Una volta individuato l'esatto intervallo temporale da utilizzare come *baseline* per ciascun soggetto, è fondamentale assicurarsi che i dati numerici registrati al suo interno siano fisicamente attendibili. Per questo motivo, prima di procedere al calcolo matematico dei valori di riferimento, il codice applica dei filtri di qualità mirati a escludere dati biologicamente incoerenti, cioè valori anomali che non riflettono in alcun modo la reale fisiologia del soggetto, ma indicano semplicemente un momentaneo distacco o un malfunzionamento dei sensori. L'algoritmo di estrazione prevede due soglie di validità specifiche:

- **Soglia per la frequenza cardiaca (MIN_HR = 50 bpm):** Il primo filtro riguarda l'attività cardiovascolare. Considerando che i soggetti monitorati sono svegli, attivi e inseriti in un contesto terapeutico, una frequenza cardiaca inferiore a 50 battiti al minuto risulta fisiologicamente irrealistica, essendo una condizione tipica solo delle fasi di sonno profondo o di atleti a riposo. Valori pari a zero o intorno ai 30 bpm rappresentano chiari artefatti dovuti alla temporanea perdita di aderenza del sensore ottico (PPG) alla pelle del polso. Dunque, se questi valori venissero inclusi nel calcolo, abbasserebbero la media drasticamente, falsando l'intera analisi. Pertanto, l'algoritmo ignora automaticamente e scarta qualsiasi dato al di sotto di questa soglia.
- **Soglia per la conduttanza cutanea (MIN_EDA = 0.05 μ S):** Il secondo filtro agisce sull'attività elettrodermica. Fisiologicamente, la pelle umana viva presenta sempre una minima conducibilità elettrica, dovuta alla naturale idratazione cutanea. Letture pari allo zero assoluto (o a valori infinitesimali come 0.01 μ S) non indicano uno stato di estremo rilassamento del paziente, ma significano semplicemente che gli elettrodi stanno misurando l'aria o la plastica perché il braccialetto si è slacciato. La scelta di posizionare la soglia di taglio esattamente a 0.05 μ S, e non a valori più elevati, è il risultato di una precisa analisi. Come osservato nel caso del Soggetto 2, esistono individui *low responder*, la cui pelle oppone un'alta resistenza e genera segnali naturali estremamente deboli, misurabili intorno agli 0.08 μ S. Impostare un filtro a 0.05 μ S permette quindi di scartare con

sicurezza il rumore del sensore staccato, garantendo al contempo di conservare e analizzare correttamente i segnali validi dei soggetti fisiologicamente poco reattivi, evitando il rischio di scartare informazioni preziose [11].

Applicati questi due filtri di sicurezza, l'algoritmo isola in modo definitivo le sessioni prese come riferimento e calcola, per ciascun soggetto, le medie matematiche della frequenza cardiaca e della conduttanza cutanea valide. La scelta di utilizzare l'operatore matematico della media non è casuale, ma risponde a una precisa necessità fisiologica e statistica. I segnali biologici, infatti, non sono mai perfettamente piatti o costanti: anche in una condizione di assoluta calma, il ritmo cardiaco accelera e decelera naturalmente assecondando la respirazione, e la pelle presenta continue e impercettibili fluttuazioni elettriche. Se si scegliesse di utilizzare come riferimento un singolo istante temporale, un valore massimo o un valore minimo, si otterrebbe un dato fragile, fortemente influenzato da un singolo respiro o dal normale rumore di fondo. Il calcolo della media, al contrario, permette di compensare tutte queste naturali micro-oscillazioni restituendo un valore stabile e altamente rappresentativo del reale stato di quiete del soggetto durante l'intera finestra temporale analizzata. Questi due precisi valori diventano così il parametro di riferimento oggettivo grazie al quale sarà possibile, nelle fasi successive dell'algoritmo, calcolare per sottrazione le variazioni assolute (ovvero i *delta*) permettendo al sistema di valutare se nei minuti successivi, il soggetto starà andando incontro a una reale attivazione da stress (registrando valori nettamente superiori alla media calcolata) o a un effettivo rilassamento (scendendo al di sotto di essa).

6.2 Finestratura temporale

Definito lo "zero clinico" per ciascun soggetto, si passa all'elaborazione del segnale continuo vero e proprio. I modelli di *Machine Learning*, tuttavia, non possono elaborare un flusso di dati temporali infinito in un'unica soluzione; richiedono invece dati discretizzati e strutturati. Per ovviare a questo problema, il codice implementa una tecnica di segmentazione del segnale nota come *windowing* (finestratura temporale). L'intero tracciato acquisito durante la sessione viene matematicamente suddiviso in

finestre o blocchi temporali più piccoli, all'interno dei quali verranno successivamente calcolati i parametri di interesse. Vengono utilizzate finestre della durata esatta di 60 secondi, seguendo le evidenze documentate negli studi [2]. Questo perché, nella ricerca sull'autismo tramite sensori indossabili, la scelta dell'ampiezza temporale rappresenta un compromesso ingegneristico e fisiologico decisivo; finestre troppo brevi (es. 5 o 10 secondi) conterebbero troppi pochi battiti cardiaci per restituire valori affidabili, rendendo i dati vulnerabili al rumore di fondo, mentre finestre troppo ampie (es. 5 minuti) non permetterebbero la rilevazione di reazioni fisiologiche o emotive rapide, come uno scatto di rabbia o un momento di paura improvvisa. La scelta dei 60 secondi rappresenta quindi il compromesso ideale: è un intervallo di tempo sufficientemente lungo per garantire dati matematicamente solidi, ma al contempo abbastanza breve da permettere all'algoritmo di rilevare variazioni fisiologiche immediate.

Per garantire la massima precisione nell'estrazione delle *feature*, l'algoritmo non esegue un taglio netto e consequenziale dei minuti (es. 0-60 secondi, 61-120 secondi), ma applica una tecnica di finestrazione scorrevole con un *overlap* di 30 secondi tra un blocco e il successivo. Di conseguenza, si analizza la prima finestra da 0 a 60 secondi, la seconda da 30 a 90 secondi, la terza da 60 a 120 secondi, e così via per l'intera durata del segnale. Questa sovrapposizione serve a non perdere informazioni cruciali nei momenti di passaggio tra un minuto e l'altro, infatti, se venissero utilizzate finestre chiuse e non sovrapposte, uno scatto d'ansia iniziato al secondo 59 finirebbe diviso a metà tra due blocchi. Matematicamente, questo significa che il valore del picco verrebbe dimezzato, risultando troppo debole in entrambe le finestre per far scattare l'allarme. Per questo motivo avviene la sovrapposizione dei blocchi temporali, proprio per evitare che questi eventi vengano diluiti, permettendo così al modello di analizzare le crisi nella loro interezza.

Infine, anche all'interno di questo ciclo di elaborazione iterativo, il codice esegue un rigoroso controllo di qualità locale, ovvero, per ogni nuova finestra di 60 secondi generata, il sistema ricalcola la media dell'EDA e verifica che essa superi la soglia di validità minima ($0.05 \mu S$) definita in fase di calibrazione. Qualora la media del blocco scenda sotto tale limite, l'algoritmo etichetta la finestra come artefatto hardware e ne forza lo scarto. In questo modo si impedisce che dati vuoti o spuri inquinino il

dataset finale, assicurando che il modello venga addestrato esclusivamente su finestre temporali biologicamente valide.

6.3 Calcolo dei parametri fisiologici

Per ogni singola finestra validata da 60 secondi, viene estratto un set specifico di metriche, trasformando in modo definitivo il segnale biologico nei numeri discreti che andranno in input al modello predittivo. La selezione di queste specifiche feature non è stabilita da criteri empirici, ma si basa sulle più recenti evidenze neuroscientifiche legate al disturbo dello spettro autistico, assegnando a ciascun parametro un ruolo clinico ben definito.

6.3.1 L'attività elettrodermica (EDA)

L'elaborazione della conduttanza cutanea e l'estrazione dei suoi parametri rappresenta la componente più complessa del sistema, agendo come vero e proprio misuratore del grado di allerta del sistema nervoso simpatico.

- **Livello tonico logaritmico** (EDA_Tonic_Log): Questo parametro definisce lo stato di base dell'attivazione. Come dimostrato da Panju et al. [11], i dati EDA grezzi presentano spesso distribuzioni distorte dovute alla crescita non lineare dell'attività tonica. Per questo motivo, il codice applica la trasformazione logaritmica $\ln(x + 1)$, formula che non solo permette di evitare errori matematici in caso di segnale nullo (poiché il logaritmo di zero tenderebbe a -infinito), ma normalizza la distribuzione. Ciò permette all'algoritmo di confrontare correttamente e sulla stessa scala i soggetti definiti *low responder* (con valori prossimi allo $0.05 \mu\text{S}$) con i soggetti *high responder*, identificando correttamente stati cronici di iper-attivazione o di inibizione da stress.
- **Pendenza tonica** (EDA_Slope): Questo parametro quantifica l'andamento generale del segnale nel minuto in analisi. Dato che i segnali biologici sono sempre frastagliati, l'algoritmo calcola questo dato tracciando una singola linea retta di tendenza (tramite interpolazione lineare) che attraversa e riassume al meglio tutti

i punti registrati in quella finestra temporale. Calcolare l'inclinazione di questa retta è considerata la feature più importante per la predizione temporale. Si fonda sul modello dell'anticipatory rise descritto da Ferguson et al. [9], secondo cui l'aggressività o la crisi non sono quasi mai precedute da un picco istantaneo, ma da un accumulo di tensione misurabile come un aumento graduale della pendenza nei minuti precedenti. Una pendenza positiva e persistente segnala quindi all'algoritmo un rischio di "esplosione emotiva", mentre una pendenza negativa indica che il sistema nervoso si sta calmando o si sta abituando allo stimolo.

- **Ampiezza e frequenza** (`Phasic_Max` e `Driver_Peaks`): Per distinguere un momento di stress puro dal sano interesse esplorativo, l'algoritmo implementa la scoperta di Greenlee et al. [10]. Il codice estrae separatamente l'altezza massima del singolo picco (`Phasic_Max`), utile a predire lo stress o l'emozione negativa improvvisa, e il numero di impulsi nervosi inviati alle ghiandole sudoripare (`Driver_Peaks`). Una densità elevata di questi ultimi, pur in presenza di un'ampiezza moderata, permette al modello di predire uno stato di engagement e di attivazione positiva, evitando di confondere l'eccitazione ludica con una crisi di rabbia.

6.3.2 Metriche cardiache e motorie

Il secondo gruppo di parametri estratti quantifica l'attività cardiovascolare e il contesto motorio, fondamentali per contestualizzare le letture elettrodermiche.

- **Media e variazione cardiaca** (`HR_Mean` e `Delta_HR`): Goodwin et al. [6] e Kushki et al. [8] evidenziano come i soggetti ASD presentino spesso una frequenza cardiaca basale cronicamente elevata. Tuttavia, il calcolo del `Delta_HR` (cioè la variazione del battito rispetto alla baseline personale) permette di catturare le accelerazioni improvvise. Un aumento tonico dell'HR senza un successivo recupero funge da indice predittivo di ansia anticipatoria o stress prolungato.
- **Variabilità della frequenza cardiaca** (`HRV_RMSSD`): Questo parametro corrisponde alla radice quadrata della media dei quadrati delle differenze successive tra i battiti. Come analizzato in fase di pre-processing, questa metrica

cattura i micro-adattamenti istantanei del ritmo cardiaco isolando l'azione del nervo vago (parasimpatico). L'RMSSD quantifica quindi l'attività del sistema nervoso parasimpatico. Come osservato nei modelli in letteratura, le crisi e i comportamenti aggressivi si verificano prevalentemente nel profilo High-High, ovvero alta conduttanza cutanea combinata a un crollo dell'HRV. Assegnando alla finestra temporale il valore RMSSD calcolato sull'intera sessione, l'algoritmo possiede il parametro chiave per comprendere se, nonostante l'eccitazione simpatica, il soggetto conserva la sua capacità di auto-regolarsi, agendo come un fondamentale fattore protettivo.

- **Dinamica del movimento (ACC_Mean):** Il contesto motorio viene quantificato calcolando la media dei valori della magnitudo vettoriale registrati dall'accelerometro (ACC_Mean), che funge da vero e proprio filtro per le letture cardiache, implementando a livello algoritmico la logica di differenziazione suggerita da Lydon et al. [27]. In ottica di classificazione, questa feature funge da filtro per disambiguare le letture cardiache, implementando la logica di differenziazione suggerita da Lydon et al. [27]. Incrociando l'attività del cuore con l'effettivo movimento corporeo, l'algoritmo discrimina matematicamente la tachicardia da sforzo fisico (esplorazione/gioco) dalla tachicardia da *distress*. Questo disaccoppiamento tra attività cardiaca e movimento permette al modello di identificare il reale sovraccarico interno, segnalando stati di forte ansia, allarme o freezing (il tipico "congelamento" motorio da paura), evitando così di confondere un semplice e non allarmante esercizio fisico con una crisi emotiva.

Al termine di questo processo estrattivo per ogni soggetto e sessione, le feature sono state tabulate strutturalmente all'interno del file `Dataset_Features_Complete.csv`, generando il dataset finale su cui si basa l'analisi successiva.

6.4 Analisi delle metriche estratte

L'estrazione delle feature ha permesso di quantificare le risposte fisiologiche dei soggetti di fronte a differenti compiti (sforzo cognitivo, induzione di stress, tecniche di

rilassamento). In primo luogo, i risultati sono stati visualizzati mediante grafici a barre (6.1 e 6.2) che illustrano le variazioni medie rispetto alla baseline. La chiave di lettura fondamentale per l'interpretazione di questi dati risiede nel concetto di variazione: la linea dello zero fisiologico rappresenta lo stato di normale riposo. Valori positivi, cioè al di sopra dello zero definito, indicano un'attivazione o eccitazione del soggetto (per esempio tachicardia o sudorazione maggiore del normale), mentre valori negativi definiscono un rilassamento e una de-attivazione rispetto alla quiete abituale.

Se da un lato questi grafici offrono un quadro clinico generale sul profilo del soggetto, dall'altro affidarsi esclusivamente alle medie matematiche dell'intera sessione presenta un limite rilevante, poiché la media tende ad appiattire le dinamiche temporali. Un valore medio elevato di conduttanza cutanea, infatti, non permette di distinguere se il soggetto è rimasto costantemente in tensione o se ha subito una serie di picchi fasici violenti alternati a crolli. Dunque, per addestrare in modo efficace il modello predittivo, è stato necessario passare da una analisi "macro" ad una "micro" per analizzare le serie temporali minuto per minuto, sfruttando le metriche estratte nel dataset completo. Confrontando le medie generali mostrate nei grafici con questo andamento fisiologico puntuale, è emerso come mente e corpo possano reagire agli stimoli in modo del tutto indipendente (registrando, ad esempio, un'alta eccitazione emotiva associata a un battito cardiaco calmo). I tre casi clinici presentati di seguito dimostrano concretamente questo concetto: solo incrociando i diversi segnali è possibile comprendere il reale stato emotivo dei soggetti e dettagli che le semplici medie non riescono a mostrare.

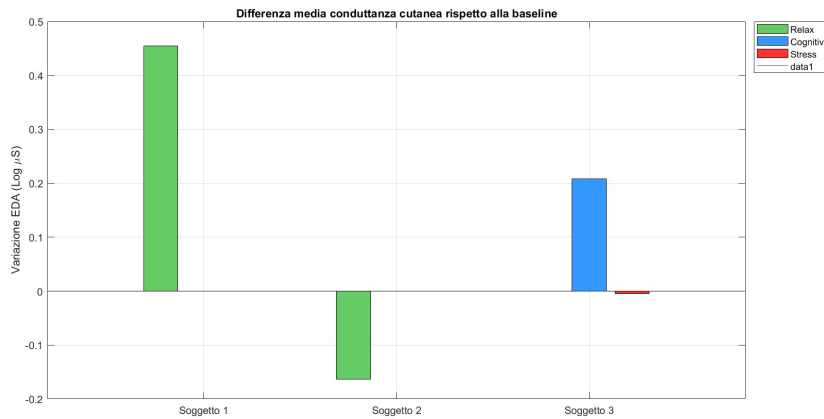


Figura 6.1: Differenza media conduttanza cutanea (EDA) rispetto alla baseline.

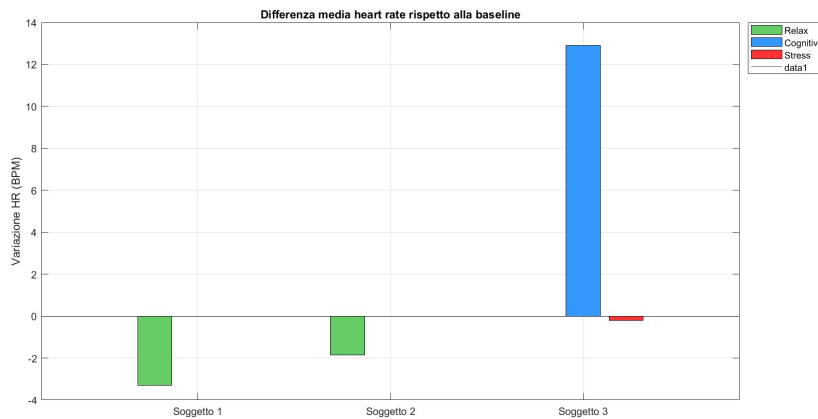


Figura 6.2: Differenza media heart rate (HR) rispetto alla baseline.

6.4.1 Soggetto 1: attivazione emotiva e capacità autoregolatoria

Il caso del Soggetto 1 si è rivelato il più complesso da decodificare. Osservando i grafici a barre durante la condizione di rilassamento, si evince una marcata contraddizione fisiologica: mentre il cuore rallenta in modo significativo (-3.5 BPM), la conduttanza cutanea aumenta (+0.45 log μ S) a causa dell'interazione sociale. L'estrazione delle *feature* traduce questa dinamica in una precisa regola di classificazione. L'elevata attivazione sudomotoria (Phasic_Max a 3.83 e ben 108 Driver_Peaks) non sfocia in aggressività perché la metrica parasimpatica HRV_RMSSD si mantiene a 97 ms. Questo specifico *pattern* numerico istruisce il modello predittivo a non etichettare automatica-

mente i picchi fasici come imminenti aggressioni (riducendo i falsi positivi), a meno che essi non siano contestualmente accompagnati da un crollo del parametro HRV.

6.4.2 Soggetto 2: il profilo low Responder e l'engagement sensoriale

Nel grafico a barre, il Soggetto 2 mostra costantemente valori negativi durante le fasi di rilassamento: la frequenza cardiaca scende (-2 BPM) così come l'attività elettrodermica (-0.15 log μ S). Questo quadro descriverebbe uno stato di rilassamento totale; ma bisogna prestare attenzione, poiché, se l'algoritmo si limitasse a osservare i soli valori assoluti minimi (EDA_Tonic_Log intorno a 0.05 μ S), rischierebbe di classificare il caso come un profilo ipo-attivo [11]. Allora, per trattare tali segnali deboli e fuorvianti, bisogna tenere conto della densità fasica; L'estrazione delle feature Drive_Peaks rivela infatti un'alta frequenza di impulsi (fino a 44 picchi al minuto), associata a un lieve abbassamento dell'HRV (43 ms) dovuto allo sforzo di concentrazione visiva. Questa combinazione di metriche insegna al modello che, nei profili *low responder*, l'assenza di variazioni macroscopiche non equivale ad apatia o *shutdown*, permettendo di identificare correttamente gli stati di sano *engagement*.

6.4.3 Soggetto 3: Reattività e coinvolgimento attivo

Come si osserva nel grafico, in risposta al compito cognitivo si registra un aumento significativo della frequenza cardiaca (quasi +13 BPM) e un parallelo incremento dell'EDA (+0.2 log μ S). Questa doppia attivazione rappresenta la classica risposta metabolica allo sforzo mentale prolungato.

Ma dalle feature estratte, confrontando la fase di induzione farmacologica (sessione 1) con la terapia di pressione profonda (sessione 2), si evince come l'effetto del farmaco si traduce matematicamente in una pendenza negativa estrema (EDA_Slope a -0.009) e nel crollo a zero del valore Phasic_Max, documentando una inibizione chimica forzata. Al contrario, durante la pressione profonda della seconda sessione, l'EDA tonica si mantiene alta, ma l'HRV resta protettivo (74 ms). Il soggetto rappresenta un profilo di basso rischio poiché si trova in uno stato di calma vigile, connesso all'am-

biente senza generare stress cardiaco. L'algoritmo predittivo impara da questo soggetto che un'elevata attività tonica non sfocia in meltdown se adeguatamente supportata da pratiche di regolazione somatica.

7. Modello predittivo Random Forest

L'obiettivo del presente lavoro è lo sviluppo di un sistema predittivo in grado di anticipare gli stati di sovraccarico fisiologico e il rischio di alterazioni comportamentali in soggetti ASD a basso funzionamento che potrebbero sfociare in aggressività. In linea con quanto effettuato da Goodwin et al. [1], il modello progettato segue una architettura *person-dependent* (soggetto-specifica). A sostegno di questa decisione, al posto di un modello globale, vi è la necessità di rispettare l'estrema eterogeneità fenotipica dell'autismo: comportamenti che aiuterebbero un soggetto a calmarsi, come ad esempio movimenti ripetitivi autoregolatori, per un altro soggetto possono indicare uno stato di forte stress e sovraccarico. Per la realizzazione di tale modello è stata cruciale l'analisi dettagliata dei dati fisiologici estratti e delle annotazioni presenti sul protocollo.

7.1 Definizione del target clinico

Per permettere a un algoritmo di apprendere, è necessario definire matematicamente la condizione di rischio (il *ground truth*) che deve essere in grado di prevedere. Anziché adottare una soglia fissa universale (es. "vi è rischio se la conduttanza supera un certo valore"), sono state applicate regole cliniche soggetto-specifiche, derivate dallo stato dell'arte e calcolate sui parametri individuali. Nel dettaglio:

- **Soggetto 1: la capacità di autoregolazione:** Il modello si fonda sulla teoria polivagale [32] e sul rischio di ritiro del freno vagale [15]. Il Soggetto 1 presenta reazioni simpatiche esplosive in risposta al cambio di contesto sociale [5]. Ai fini della predizione, queste risposte di orientamento sociale sono tollerate fintanto che il sistema parasimpatico le compensa, dunque l'allarme scatta esclusivamente in presenza di un profilo che presenta elevata eccitazione EDA combinata a un crollo dell'HRV.
- **Soggetto 2: il paradosso dello stimming:** per questo soggetto *low responder*, è stata implementata la teoria del "paradosso dell'oggetto" [12] e la teoria dell'in-

terferenza [4]. Il Soggetto 2 utilizza piccoli movimenti del corpo per processare gli stimoli e autoregolarsi. Pertanto, i picchi di conduttanza cutanea non vengono penalizzati se supportati da un corrispondente movimento. Il rischio viene rilevato dall'algoritmo solo quando si verifica un blocco motorio anomalo: se a un'alta attivazione nervosa corrisponde un arresto dei movimenti, il comportamento rassicurante del soggetto è stato interrotto, generando possibile frustrazione e crisi.

- **Soggetto 3: il sovraccarico emotivo graduale:** basato sull'evidenza dell'*anticipatory rise* [9], il rischio per questo soggetto non è dato da un singolo picco improvviso, ma dall'accumulo di tensione. Per addestrare il modello a riconoscere questo specifico fenomeno, è stata applicata la seguente distinzione: mentre nel capitolo precedente la fase di recupero post-farmaco è servita a calibrare il livello base dei sensori, in questa fase predittiva la sessione di sforzo cognitivo positivo è stata utilizzata per definire il limite massimo della zona di sicurezza; in questo modo l'algoritmo impara a non confondere l'impegno cognitivo sano con l'inizio di una crisi. L'allarme scatta, infatti, solo se i picchi di sudorazione superano questi livelli di normale concentrazione e rappresentano il culmine di una tensione che sale in modo rapido e persistente, annunciando un imminente attacco di panico.

7.1.1 Definizione delle soglie di rischio in assenza di crisi

L'impostazione delle regole cliniche appena descritte è stata fondamentale per tracciare la logica del sistema. Tuttavia, analizzando i tracciati, è emerso che nelle sessioni a disposizione non si sono verificati episodi di aggressività visibile. Un'applicazione rigida delle sole regole cliniche avrebbe di fatto impedito l'addestramento del modello, per la semplice mancanza di eventi critici esterni. Invece di considerare questo aspetto come un limite ai fini dell'addestramento, l'approccio è stato calibrato per sfruttare la vera potenzialità dei sensori biometrici: rilevare la sofferenza invisibile.

In assenza di crisi esplosive, l'algoritmo è stato istruito a etichettare come target di rischio il 10% dei picchi fasici massimi assoluti della giornata. Pur senza sfociare in

un *meltdown* comportamentale evidente, quel 10% di picchi rappresenta il momento di massimo sforzo, accumulo di tensione e *iper-arousal* silente per il sistema nervoso del soggetto in quel giorno. In questo modo, il modello non si limita a diagnosticare una crisi già visibile a occhio nudo, ma si addestra a riconoscere e prevedere i momenti di massimo sovraccarico interno prima che la situazione degeneri. Questo rappresenta l'esatto scopo clinico e preventivo dell'utilizzo dei sensori indossabili nell'autismo.

Le soglie numeriche utilizzate per definire queste anomalie derivano dalle regole dell'*anomaly detection* applicate all'elaborazione dei segnali biomedici [33]. I segnali fisiologici naturalistici sono infatti intrinsecamente rumorosi; pertanto, per distinguere un reale evento di stress clinico dalle normali fluttuazioni del corpo, sono stati adottati due approcci statistici distinti:

- **L'uso dei percentili per l'attivazione simpatica (85° e 95°):** L'attività elettrodermica è fortemente influenzata dalla normale termoregolazione del corpo [29]. Se si fosse scelta una soglia assoluta, si avrebbe rischiato di confondere il normale sudore per il caldo con una crisi d'ansia. Impostando la soglia all'85° o al 95° percentile, l'algoritmo isola esclusivamente il 15% o il 5% dei picchi più alti registrati dal soggetto. In statistica, questi percentili approssimano il limite di controllo situato oltre $+1\sigma$ e $+1.64\sigma$ dalla media [34](Figura 7.1).
- **L'uso delle mediane e delle soglie di tolleranza per il crollo motorio e vagale:** Quando si analizza il movimento o l'HRV, l'uso della classica media matematica è sconsigliato, poiché un singolo movimento brusco o un artefatto del sensore la falserebbe. Per questo motivo è stata utilizzata la mediana, che rappresenta il vero valore centrale e stabile di riferimento del soggetto, immune agli sbalzi isolati [35]. Tuttavia, stabilito questo valore centrale, bisognava decidere di quanto dovesse scendere il segnale per far scattare l'allarme. Sono state applicate soglie dell'80% per il movimento (cioè un calo del 20%) e del 90% per l'HRV (un calo del 10%). Queste specifiche percentuali rispondono al criterio clinico del *minimum detectable change* [36], ovvero il minimo cambiamento Rilevabile. I sensori da polso presentano sempre una minima percentuale di errore e la frequenza cardiaca oscilla naturalmente con il respiro [36]. Richiedendo un crollo

netto del 10% o del 20%, ci si assicura di superare questo margine di errore naturale: l'algoritmo non lancia l'allarme per un micromovimento del braccialetto, ma certifica matematicamente che è in atto un reale blocco motorio o un ritiro del freno vagale.

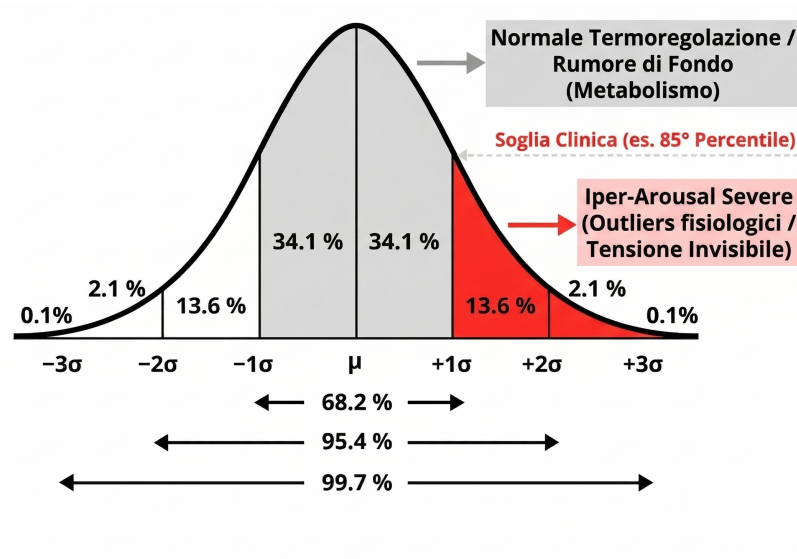


Figura 7.1: Schema concettuale del rilevamento anomalie nell'EDA. L'area grigia rappresenta la normale varianza metabolica, mentre l'area rossa evidenzia gli outlier fisiologici (iper-arousal) isolati tramite i percentili.

7.2 La finestra di previsione e il tempo di intervento

Se l'algoritmo facesse scattare l'allarme solo nel momento esatto in cui la crisi raggiunge il suo apice, sarebbe un sistema diagnostico, non predittivo. Per trasformare il braccialetto in un vero strumento di prevenzione, è stata adottata una tecnica chiamata *label shifting*, consistente nella traslazione del target. In altre parole, si è progettato l'algoritmo non per rilevare il picco della crisi, ma per tornare indietro di 3 finestre temporali (da 30 secondi ciascuna), ovvero di 1.5 minuti.

In questo modo, il modello viene matematicamente forzato a studiare gli indicatori che nel corpo del soggetto precedono il sovraccarico. La scelta di questo intervallo di 90 secondi non è casuale, ma si allinea all'orizzonte predittivo validato in letteratura da

Goodwin et al. [1]. Questo lasso temporale deriva infatti da un preciso compromesso tra la biologia del corpo e il tempo pratico di intervento. Infatti, questo intervallo rispetta il tempismo fisiologico corporeo, dal momento che i cambiamenti ormonali e la reazione delle ghiandole sudoripare non avvengono istantaneamente ma hanno precisi tempi di latenza [37]. Inoltre, un minuto e mezzo è il tempo vitale necessario a un educatore o a un genitore per accorgersi del pericolo imminente e intervenire con strategie calmanti, evitando che la tensione degeneri in aggressività.

A livello matematico e algoritmico, questa operazione si traduce in una mappatura dei parametri fisiologici acquisiti al tempo T verso un evento target al tempo $T + \Delta t$ (con $\Delta t = 1.5$ minuti). Nello specifico dell'implementazione del codice, se il sistema identifica la crisi all'istante discreto k , l'etichetta di rischio clinico (il target per l'addestramento, Y_{train}) non viene limitata a quel singolo indice, ma viene retropropagata. L'algoritmo istruisce la cosiddetta fase di allarme assegnando valore logico positivo (*true*) a tutte le finestre temporali comprese nell'intervallo $[\max(1, k - n), k]$, dove n rappresenta il numero di passi di anticipo temporale impostati ($n = 3$).

In questo modo, il Random Forest non si limita a diagnosticare la crisi già avvenuta, ma impara a riconoscere in anticipo i pattern fisiologici precursori che portano al sovraccarico. Infine, a livello logico, l'utilizzo della funzione `max` serve a proteggere il codice da eventuali blocchi. Se infatti un evento critico dovesse verificarsi nei primissimi minuti di utilizzo del braccialetto, l'algoritmo cercherebbe di tornare indietro nel tempo a un momento in cui il sensore non stava ancora registrando, generando un errore matematico (ovvero indici negativi). La funzione `max` impedisce che questo accada, assicurando che il sistema non vada mai 'sotto zero' e che il calcolo parta sempre dal primissimo campione valido a disposizione.

7.3 Architettura, addestramento e validazione del modello

Per l'elaborazione dei vettori di feature e la generazione della probabilità di rischio futuro, è stato impiegato il classificatore Random Forest. La scelta di un algoritmo di *ensemble learning* è fortemente supportata dalla letteratura sull'*affective computing*

applicata all'elaborazione dei dati fisiologici provenienti dai *wearable* [39]. Questo approccio si basa sulla logica di non affidare la predizione a un unico modello isolato, ma di addestrare molteplici alberi decisionali, in cui l'output finale emerge da una votazione a maggioranza capace di compensare le imprecisioni dei singoli nodi. Questa architettura robusta è necessaria perché i segnali corporei, in particolar modo l'attività elettrodermica (EDA) e la variabilità cardiaca (HRV), presentano una varianza elevata e sono intrinsecamente non lineari. Il sistema nervoso umano, infatti, non risponde agli stimoli ambientali in modo meccanico o proporzionale: l'intensità di una reazione fisiologica dipende fortemente dallo stato di attivazione di partenza del soggetto e i vari parametri si influenzano a vicenda creando dinamiche complesse che una semplice retta matematica non può descrivere. Il Random Forest, proprio grazie alla sua struttura ramificata a bivi successivi, riesce ad assecondare e decodificare questa natura dinamica della biologia umana, garantendo prestazioni predittive superiori rispetto ai tradizionali modelli lineari.

Dunque, anziché generare una singola regola decisionale che rischierebbe di rivelarsi troppo vincolante, l'algoritmo è stato configurato per costruire una vera e propria "foresta" di 50 alberi decisionali indipendenti (impostando il parametro `NumLearningCycles = 50`). Per ottenere questa indipendenza, è stata applicata la tecnica del *bootstrap aggregating* o *bagging* [38]: in pratica, durante l'addestramento, non è stato fornito a tutti gli alberi lo stesso blocco di informazioni. Al contrario, ogni albero ha studiato un sottoinsieme casuale e diverso dei dati, venendo costretto a cercare in totale autonomia le proprie regole matematiche per separare i minuti di "calma" da quelli di "rischio". Successivamente, in fase di predizione, i nuovi dati fisiologici del soggetto vengono sottoposti simultaneamente a tutti i 50 alberi. In questo modo, la curva predittiva tracciata dal sistema non è un semplice output binario, ma è il risultato di una *soft classification*. Questo significa che, anziché imporre una rigida classificazione finale di "crisi" o "calma", il sistema traduce le decisioni della "foresta" di alberi in una probabilità fluida e continua. Il livello di rischio che si osserva salire nel grafico rappresenta infatti l'esatta frazione dei voti positivi dell'intero *ensemble*: se 40 alberi su 50 riconoscono i precursori di un sovraccarico, il modello restituirà un segnale di allarme progressivo ed esatto dell'80%. Questo meccanismo a maggioranza permette

agli alberi di correggersi a vicenda; infatti, se un singolo albero viene ingannato da un rumore isolato, come un banale artefatto da movimento, il resto della foresta compensa l'errore. In questo modo il sistema diventa solidissimo e azzerà i falsi allarmi.

Per dimostrare la validità scientifica della predizione ed evitare il rischio di *overfitting*, ovvero l'errore per cui l'algoritmo si limita a imparare a memoria i dati di addestramento perdendo la capacità di adattarsi a situazioni nuove, le performance dei modelli sono state valutate attraverso una rigorosa *k-fold cross-validation* a 5 vie. Nella pratica, il dataset di ogni paziente è stato suddiviso casualmente in 5 blocchi distinti (*fold*). L'algoritmo è stato addestrato iterativamente su 4 di questi blocchi ed è stato poi obbligato a testare le proprie previsioni sul quinto, un blocco di dati rigorosamente tenuto nascosto. Questa procedura è stata ripetuta 5 volte, ruotando di volta in volta il blocco di test. In questo modo l'algoritmo è stato messo alla prova solo su dati "nuovi", mai incontrati durante l'addestramento.

Successivamente, per quantificare l'efficacia del sistema in ambito clinico, sono state estratte tre metriche fondamentali: l'accuratezza misura la correttezza globale delle previsioni; la sensibilità rappresenta la capacità vitale del sistema di intercettare i reali pattern di rischio, garantendo che nessuna potenziale crisi venga mancata; la specificità, al contrario, misura l'abilità dell'algoritmo di riconoscere correttamente le fasi di calma fisiologica, un parametro cruciale per non sovraccaricare i caregiver con continui e inutili falsi allarmi. Questo approccio è fondamentale per evitare il *data leakage*, ovvero il rischio che i dati usati per l'addestramento e quelli usati per il test si mescolino. Poiché i risultati finali sono la media di 5 test effettuati su dati sempre diversi, si ha la prova matematica che il modello non ha semplicemente memorizzato i dati storici, ma ha imparato a generalizzare i veri precursori fisiologici delle crisi, confermandosi affidabile per l'uso clinico reale.

Infine, in ambito clinico e medico, non basta che l'algoritmo fornisca una previsione corretta ma bisogna poter capire il motivo di quella decisione; dunque, una volta superati i test di validazione, si è proceduto ad addestrare l'algoritmo definitivo sull'intero dataset per potere avere dei dettagli riguardanti la "giustificazione" delle scelte prese. A questo scopo è stata estratta la *feature importance* utilizzando una metrica matematica chiamata "indice di Gini". In parole semplici, questo indice misura la capacità

di ogni singolo sensore di fare chiarezza: ovvero calcola matematicamente quanto ogni singolo parametro abbia aiutato il modello a ridurre l'incertezza e prendere una decisione netta; ad esempio, se l'uso dell'EDA ha permesso agli alberi di separare in modo netto i momenti di calma da quelli di crisi, il suo Indice di Gini sarà altissimo. Questa metrica è la prova che l'algoritmo non agisce casualmente, ma ha imparato in totale autonomia a dare maggior peso a quei sensori che in letteratura medica sono considerati i veri indicatori del sovraccarico emotivo.

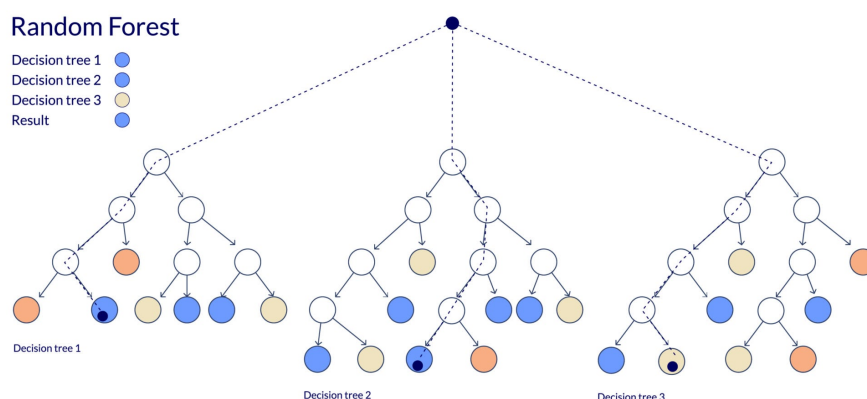


Figura 7.2: Rappresentazione schematica dell'architettura Random Forest. L'algoritmo aggrega le predizioni di molteplici alberi decisionali indipendenti per generare un risultato finale robusto e bilanciato.

7.4 Analisi dei risultati

7.4.1 Metriche di performance globale

Come mostrato nella tabella seguente, il modello presenta dei valori di precisione molto elevati:

Tabella 7.1: Metriche di performance del Random Forest (cross-validation a 5-fold)

Soggetto	Accuratezza	Sensibilità (TPR)	Specificità (TNR)	Area ROC (AUC)
Soggetto 1	95.35%	83.33%	98.53%	0.98
Soggetto 2	95.21%	88.89%	96.64%	0.97
Soggetto 3	99.49%	85.71%	100.00%	1.00

Come evidenziato dalle metriche, l'approccio *person-dependent* risulta estremamente bilanciato. Una specificità media superiore al 98% indica una pressoché totale

assenza di falsi allarmi durante le fasi di reale rilassamento o *engagement*. In egual modo, la sensibilità assicura che oltre l'85% dei pattern precursori del rischio venga intercettato in modo tempestivo, fornendo una finestra di intervento utile ai caregiver.

L'Area sotto la curva ROC (AUC), con valori compresi tra 0.97 e 1.00, rappresenta la prova matematica della capacità discriminante del sistema tra lo stato di calma e quello di rischio pre-crisi. Bisogna però osservare il valore di AUC pari a 1.00 ottenuto per il Soggetto 3. Un punteggio perfetto risulta sospetto, poiché nel *machine learning* applicato alla biologia ci si aspetta sempre un certo margine di rumore che sporca i dati; tuttavia, questo risultato trova una duplice spiegazione. Da un lato, sebbene in genere un punteggio perfetto suggerisca che il test sia stato 'troppo semplice', per questo soggetto i segnali che precedono il sovraccarico sono talmente chiari e diversi dal normale stato di riposo da non lasciare all'algoritmo alcuno spazio per le incertezze. Dall'altro lato, è doveroso riconoscere che questo risultato anomalo è strettamente connesso ai limiti attuali del dataset, caratterizzato da un ridotto numero di campioni e dall'assenza, nelle registrazioni reali, di crisi conclamate. Per questo motivo, il lavoro prevede l'integrazione di dati artificiali che permetterà di simulare scenari più complessi e realistici, aggiungendo ai dati quella variabilità che oggi manca. In questo modo si potrà testare il sistema in condizioni più 'difficili', garantendo che rimanga affidabile anche nelle situazioni imprevedibili del mondo reale.

7.4.2 Feature importance

Per valutarne la capacità decisionale, il modello è stato analizzato tramite l'indice di Gini per estrarne la *Feature Importance*, metrica che svela in modo oggettivo quali parametri fisiologici sono stati più influenti per guidare l'algoritmo verso la decisione.

Come si evince dai grafici (Figura 7.4), in tutti e tre i soggetti i parametri legati all'attività fasica (*Phasic_AUC* e *Phasic_Max*) hanno il maggior peso ai fini della predizione convalidando l'uso della risposta elettrodermica rapida come marcatore primario dell'attivazione simpatica istantanea. In particolare, nei profili del Soggetto 1 e del Soggetto 2, emerge il peso determinante del parametro *EDA_Slope*, ovvero la pendenza tonica, confermando che l'algoritmo non si è limitato a valutare l'intensità assoluta del segnale in un singolo momento, ma ha imparato ad analizzarne la derivata tempo-

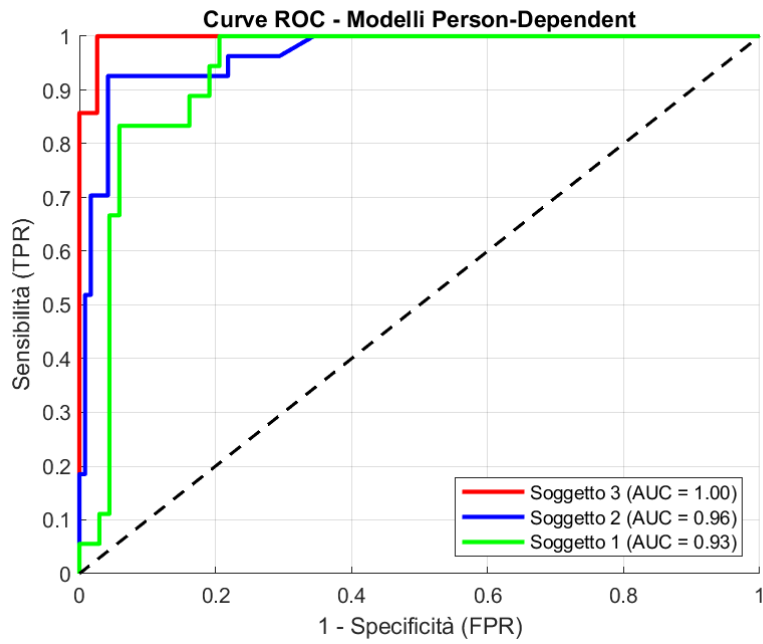
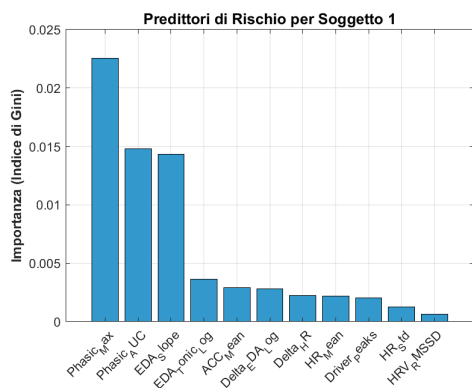
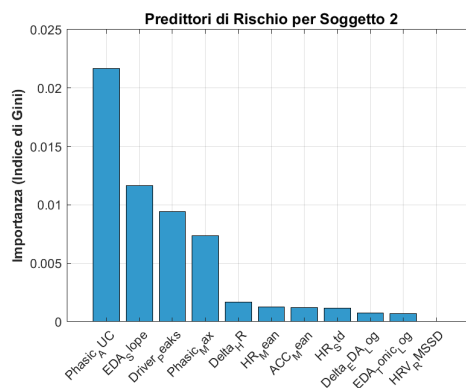


Figura 7.3: Curve ROC dei modelli person-dependent.

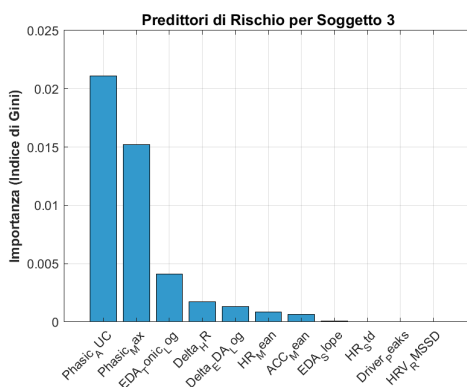
rare, ovvero la traiettoria con cui lo stress cumulativo sta crescendo. Per il Soggetto 3, invece, assume rilevanza anche il livello tonico di base (EDA_Tonic_Log), evidenziando ulteriormente come l'algorithmo implementato abbia definito le proprie regole decisionali rispettando le specifiche peculiarità fisiologiche di ogni singolo individuo.



(a) Soggetto 1



(b) Soggetto 2



(c) Soggetto 3

Figura 7.4: Estrazione della feature importance tramite Indice di Gini. Si osserva la predominanza dei parametri fasici per tutti i soggetti, affiancata dall'importanza dell'EDA Slope per i Soggetti 1 e 2, e del livello tonico di base per il Soggetto 3.

7.4.3 Il tempo di intervento sui casi reali

Per dimostrare la reale efficacia del modello, le sue previsioni sono state confrontate direttamente con gli eventi definiti nel protocollo. I grafici mostrano chiaramente il risultato pratico dell'algorithm: la creazione di una finestra di intervento, ovvero il *lead time*. Ottenere un *lead time* significa calcolare con quanti minuti di anticipo il sistema lancia l'allarme, fornendo al caregiver il tempo materiale per intervenire e bloccare la disregolazione emotiva prima che esploda in una crisi acuta. Per questa fase di validazione visiva sono state selezionate intenzionalmente le prime sessioni di registrazione di ciascun soggetto, poiché, come analizzato in precedenza, rappresentavano i contesti di maggiore vulnerabilità, in cui si sono verificati possibili episodi di sovraccarico.

Per agevolare la comprensione delle dinamiche, è utile chiarire la struttura dei grafici che accompagnano ogni caso clinico. Ciascuna figura è divisa in due pannelli sovrapposti che condividono la stessa linea del tempo (espressa in minuti sull'asse orizzontale), permettendo di confrontare istante per istante la biologia del soggetto con le decisioni dell'algorithm:

- **Pannello A (fisiologia reale):** Il grafico superiore traccia in blu l'andamento dell'attivazione fasica dell'EDA. La scelta di visualizzare proprio questo segnale non è casuale: come dimostrato nella precedente analisi della *feature importance*, la risposta elettrodermica rapida si è confermata matematicamente come il predittore più potente per tutti i soggetti. Questa curva rappresenta la reale eccitazione del sistema nervoso autonomo del soggetto. La linea verticale nera, sormontata da un punto rosso ("Evento Clinico Max"), indica il momento esatto in cui è stato annotato l'evento critico di riferimento.
- **Pannello B (risposta dell'algorithm):** Il grafico inferiore traccia in rosso la probabilità di rischio di un imminente sovraccarico, calcolata in tempo reale dall'algorithm (su una scala da 0 a 100%). All'interno del grafico è tracciata una linea orizzontale tratteggiata: questa è la "soglia di rischio", impostata al 25%. Abbassare la soglia rispetto al normale 50% permette al sistema di lanciare l'allarme non appena percepisce le prime lievi anomalie fisiologiche. In questo mo-

do si massimizza il tempo di anticipo a disposizione del caregiver, mantenendo comunque un'altissima precisione e azzerando i falsi allarmi.

- **La Finestra di intervento (area gialla):** Nel momento in cui la linea rossa della probabilità supera la soglia del 25%, il sistema fa scattare l'allarme preventivo. L'area evidenziata in giallo rappresenta visivamente proprio il *lead time*: è il tempo che intercorre tra l'istante in cui l'algoritmo lancia l'allarme e il momento in cui l'evento critico si manifesta davvero.

Di seguito vengono analizzati nel dettaglio i tre scenari clinici:

Soggetto 1: Il profilo social-reattivo

Viene documentata al minuto 16.5 un netto cambio di postura accompagnato da un'interazione sociale che provoca nel soggetto un picco di attivazione simpatica. Analizzando la curva di predizione, si osserva come l'algoritmo riesca a intercettare i precursori silenti, in particolare il crollo del freno vagale, superando la soglia di rischio del 25% già al minuto 15.0. Questo genera un *lead time* di 1.5 minuti. Dunque, il sistema riesce a cogliere l'impatto di uno stimolo sociale quando ancora non c'è alcun segnale visibile all'esterno offrendo un certo margine di tempo per intervenire e prevenire una possibile crisi innescata da tale cambio di contesto.

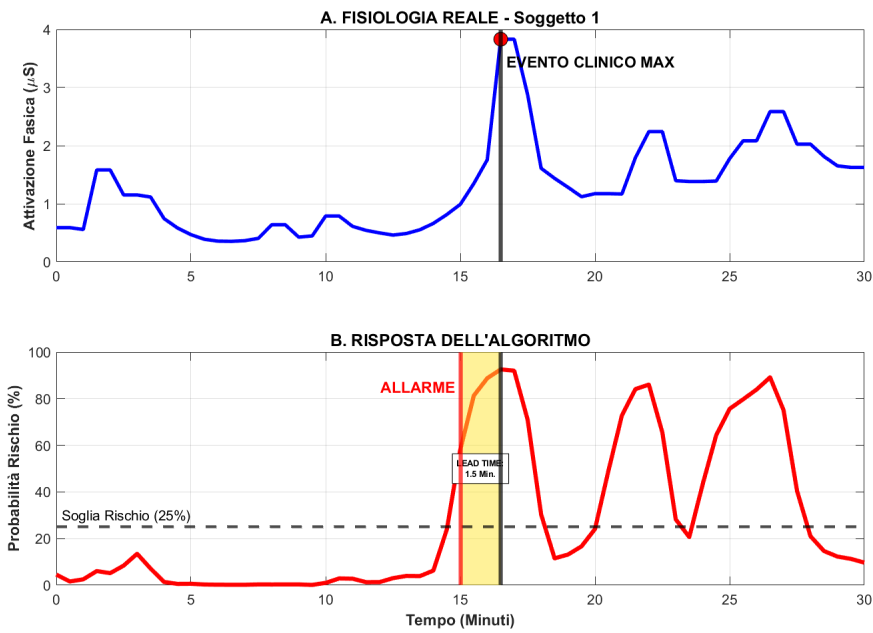


Figura 7.5: Analisi temporale e *Lead Time* per il Soggetto 1.

Soggetto 2: Il profilo sensoriale-motorio

Il soggetto raggiunge un apice di eccitazione interagendo con alcuni oggetti luminosi. In questo caso, il modello anticipa i precursori fisiologici con un notevole *lead time* di 3.0 minuti. L'aspetto più rilevante di questa analisi risiede nella fase successiva al raggiungimento dell'apice in cui il soggetto inizia a usare lo *stimming* motorio per sfogare e autoregolare la gioia; a conferma di ciò l'attività elettrodermica si mantiene su valori altissimi. Un modello lineare elementare avrebbe generato un falso allarme continuo basandosi solo sull'eccesso di sudorazione; il Random Forest, invece, fa crollare rapidamente la probabilità di rischio dal momento che incrocia i dati dell'eccitazione con quelli del movimento e, riconoscendo un movimento ritmico, capisce che l'energia in eccesso viene scaricata in modo sicuro. Il sistema valuta quindi la situazione come sotto controllo, evitando falsi allarmi che andrebbero a interrompere lo sfogo naturale della persona.

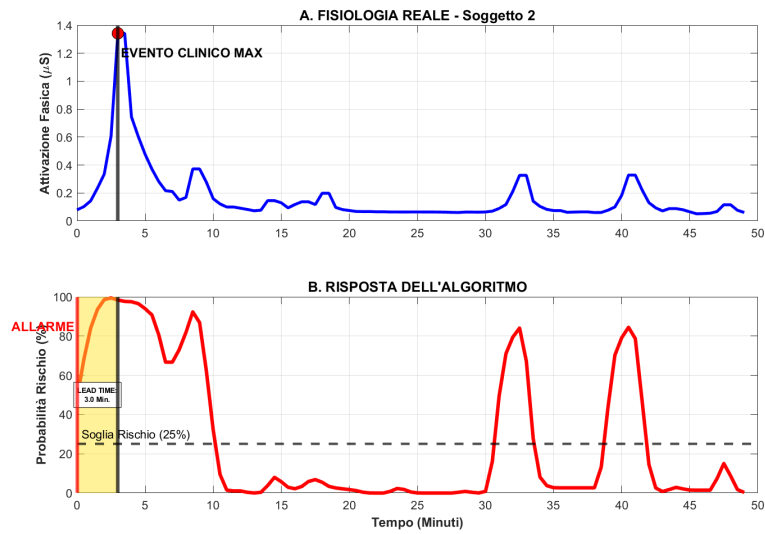


Figura 7.6: Analisi temporale e *Lead Time* per il Soggetto 2.

Soggetto 3: La validazione farmacologica

In questo caso, l'evento clinico (un attacco di panico acuto) era antecedente all'avvio della registrazione, portando logicamente l'algoritmo a calcolare un rischio predittivo iniziale del 100%. Osservando il decadimento della curva si nota che, in perfetta sincronia con i tempi di assorbimento del farmaco sedativo (*Tavor*) somministrato, la curva di rischio calcolata dall'algoritmo collassa e si azzerava. L'algoritmo ha registrato l'assenza di stress indotta dal farmaco, certificando in modo oggettivo il ritorno alla sicurezza (0% di rischio). Questa è la prova che il modello non calcola percentuali astratte, ma traccia istante per istante le reali reazioni del corpo.

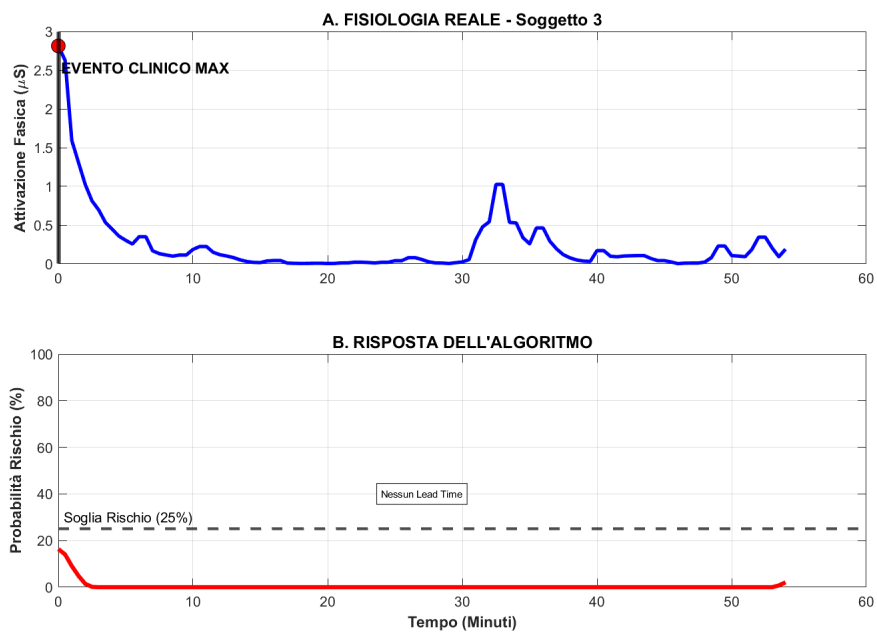


Figura 7.7: Analisi temporale per il Soggetto 3.

7.4.4 Il confronto con l'approccio non supervisionato di Change Point Analysis

I risultati ottenuti trovano conferma nel lavoro parallelo svolto all'interno del medesimo progetto dalla collega Alessia Tudino, la quale ha indagato lo stesso set di dati reali mediante un approccio non supervisionato, basato sulla *Change Point Analysis* (CPA). Il presente lavoro ha impiegato un classificatore supervisionato *Random Forest* con soglie dinamiche basate sui percentili, mentre l'approccio CPA si è focalizzato sull'individuazione delle rotture strutturali nelle medie mobili del segnale. Nonostante le differenze algoritmiche, sono emersi parallelismi e analisi complementari. Entrambi gli approcci avevano come obiettivo quello di intercettare quello stato di stress non osservabile dall'esterno. Dunque, sia il modello ad alberi decisionali che l'algoritmo basato sulla CPA sono riusciti a isolare i reali segnali di stress dal normale rumore di fondo del corpo. Questa complementarità emerge confrontando la risposta dei due algoritmi sugli stessi scenari clinici:

- **Nel caso del Soggetto 1**, durante il cambio di contesto sociale, l'analisi CPA ha individuato una deviazione improvvisa nell'attività elettrodermica esattamente nel minuto in cui è avvenuta l'interazione. Questo dato conferma in modo speculare il *lead time* di 1.5 minuti calcolato dal *Random Forest* prima dell'iperattivazione totale.
- **Per quanto riguarda il Soggetto 2**, impegnato in un'attività sensoriale stimolante, l'algoritmo CPA è stato in grado di isolare i salti di regime legati all'interazione cognitiva senza confonderli con un pericoloso accumulo di tensione. Lo stesso è accaduto nel caso del *Random Forest*, che, riconoscendo la dinamica di autoregolazione tramite il movimento ritmico, ha mantenuto bassa la probabilità di rischio; il sistema ha dunque evitato in entrambi gli approcci di generare falsi allarmi.
- **Allo stesso modo, nella validazione farmacologica del Soggetto 3**, la *Change Point Analysis* ha rilevato l'importante salto di regime indotto dal farmaco sedati-

vo, allineandosi perfettamente con il crollo istantaneo della probabilità di rischio (allo 0%) del modello ad alberi decisionali.

Inoltre, ambo le analisi hanno confermato matematicamente la stessa dinamica: l'aumento dell'attività elettrodermica precede sempre l'accelerazione cardiaca o i movimenti del corpo. In altre parole, che si usi la probabilità di una "foresta" di alberi o che si cerchi un'improvvisa variazione nella *baseline*, la conclusione clinica è la stessa: il corpo lancia segnali d'allarme silenziosi ma misurabili, molto prima che la crisi esploda e diventi visibile all'esterno.

7.4.5 Conclusioni e utilità clinica

L'integrazione dei risultati fin qui esposti dimostra che la predizione delle crisi nell'A-SD non può essere ridotta al semplice superamento statistico di una soglia di attivazione prestabilita. Il collegamento tra l'estrazione dei dati fisiologici e la previsione anticipata dell'algoritmo ha trasformato ciò che prima era solo un'intuizione clinica in solide regole matematiche.

Nel quotidiano, questo sistema garantisce le seguenti utilità cliniche:

- permette di passare dalla pura reazione all'evento alla sua reale prevenzione, fornendo minuti importanti per agire;
- rende oggettivamente visibile il sovraccarico interno, un fattore cruciale per proteggere i soggetti non verbali;
- rispetta profondamente l'eterogeneità fenotipica di questi individui grazie alla creazione di profili *person-dependent* che non penalizzano l'autostimolazione benigna;
- e, infine, offre uno strumento oggettivo per la validazione di terapie o degli ambienti di rilassamento.

Resta però un limite oggettivo nei dati reali raccolti durante il protocollo: per fortuna, non si sono verificate crisi comportamentali gravi; tuttavia, il modello ha imparato a riconoscere il sovraccarico positivo, l'attenzione e i primissimi segnali di allarme

invisibili, ma per funzionare nel mondo reale deve saper gestire anche le situazioni estreme. Dunque, per superare questa mancanza di dati, il passo successivo di questo lavoro è lo sviluppo di un modello ibrido con la generazione di dati artificiali in modo controllato, così da poter simulare la progressiva attivazione fisiologica verso una vera e propria crisi, andando a coprire ciò che manca nelle registrazioni sul campo e rendendo così il sistema molto più solido e affidabile.

8. Modello predittivo ibrido e generazione di dati sintetici

Per superare i limiti dei dati reali, l'evoluzione del lavoro di tesi illustrata nel presente capitolo si è sviluppata attraverso lo sviluppo di un modello ibrido. Tale architettura nasce dalla sinergia tra i contributi di ricerca sviluppati in parallelo: la *Change Point Analysis* (CPA) curata dalla collega Alessia Tudino e l'approccio *Random Forest* fulcro del presente studio. Nelle sezioni seguenti verrà descritto il processo matematico di generazione dei dati fisiologici sintetici e la successiva implementazione di un'architettura predittiva a due stadi. Questo nuovo approccio integra la capacità di rilevamento anomalie della *Change Point Analysis* con la solidità decisionale del classificatore *Random Forest*, al fine di garantire un sistema di allerta robusto e applicabile nel mondo reale.

8.1 L'architettura a due stadi

Come ampiamente discusso, l'analisi del dataset reale ha dimostrato l'efficacia degli algoritmi nel rilevare gli stati di attivazione fisiologica e il sovraccarico emotivo. Tuttavia, è emerso un limite strutturale legato alla natura stessa dei dati a disposizione, ovvero l'assenza di episodi di aggressività o crisi. Per avere una utilità clinica e quotidiana, il sistema deve essere esposto a scenari critici di altissima intensità, imparando a riconoscere i precursori di una crisi prima che questa degeneri [3].

Per questo motivo, l'architettura del modello ibrido è stata progettata per sfruttare i punti di forza complementari di due metodologie sviluppate in parallelo, strutturandole in due stadi a cascata:

- **Primo stadio non supervisionato:** analizza il segnale dell'attività elettrodermica alla ricerca di una rottura improvvisa rispetto alla normale varianza di fondo (la baseline). Una volta individuato il punto di rottura, il sistema, prima di far scattare l'allarme, applica una rigorosa validazione clinica incrociando i dati di sudorazione, battito cardiaco e movimento.

- **Secondo stadio supervisionato:** solo i pattern che superano questa validazione clinica vengono passati al Random Forest. L’algoritmo, precedentemente addestrato, valuta l’intera dinamica e restituisce una probabilità continua di rischio, garantendo una finestra di anticipo essenziale per l’intervento preventivo.

In altre parole, il modello esegue una logica a cascata basata su un primario rilevamento del cambiamento fisiologico seguito da una classificazione supervisionata ai fini predittivi. Questa logica adottata trova conferme nella letteratura recente; come dimostrato da Apostol et al. [19] e da Muke e Kozierekiewicz [18], l’integrazione di un rilevatore di anomalie non supervisionato prima della classificazione permette di eliminare i falsi positivi nell’analisi di segnali fisiologici multimodali, garantendo che il modello supervisionato non confonda il segnale utile con le normali fluttuazioni del corpo.

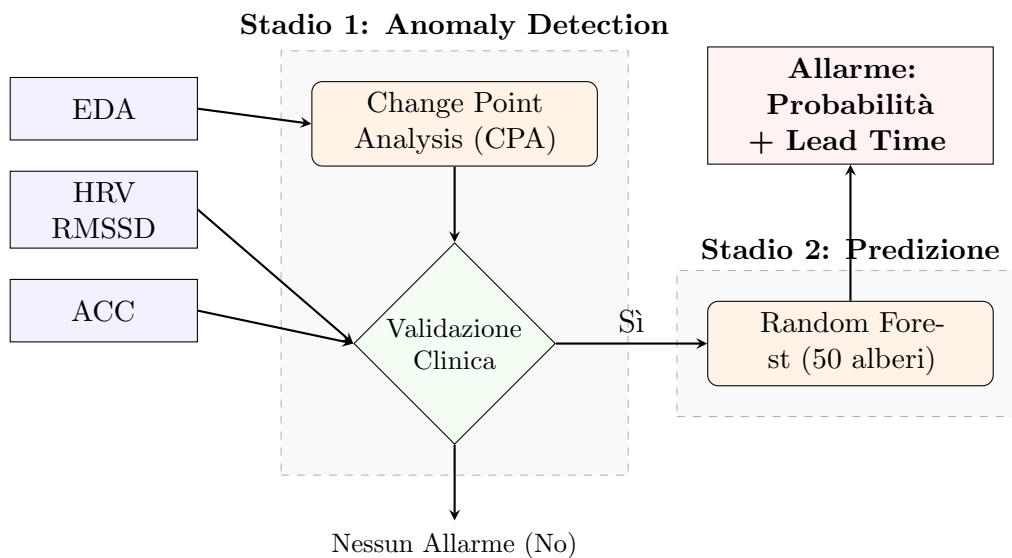


Figura 8.1: Schema del flusso logico dell’architettura ibrida a due stadi.

8.2 La generazione dei dati sintetici

Per simulare scenari di vera emergenza su cui testare questa architettura, si è ritenuto necessario ricorrere alla generazione di dati sintetici. Nel campo dell’elaborazione dei segnali biomedici, l’uso di dati artificiali rappresenta una prassi abbastanza consolidata per validare la robustezza degli algoritmi in assenza di dataset reali sufficientemente

bilanciati, un problema concreto confermato anche da Goodwin et al. [3]. Dal momento che registrare crisi comportamentali autentiche in numero sufficiente è complesso, i dati simulati diventano fondamentali per addestrare i sistemi.

A questa necessità si aggiungono le difficoltà cliniche legate alla raccolta dati sul campo: lavorare con soggetti ASD comporta limitazioni dovute a difficoltà di acquisizione (il soggetto, durante una crisi, può manifestare fastidio verso il sensore danneggiandolo o rimuovendolo), questioni etiche e di privacy stringenti relative al trattamento di dati sensibili. Inoltre, è fondamentale non indurre stress al soggetto al solo scopo di ricerca; infine, l'imprevedibilità comportamentale rende quasi impossibile pianificare sessioni di monitoraggio che garantiscano la cattura di un evento critico, portando a lunghi tempi di registrazione con pochissimi eventi utili.

Dunque, a supporto di questa scelta, Fawaz et al. [20] hanno dimostrato che la *data augmentation* migliora nettamente la precisione dei modelli su eventi rari. Più recentemente, Lange et al. [21] hanno provato che addestrare i sistemi su dati sintetici multimodali (EDA, HR, ACC) aumenta le performance fino al 15% rispetto all'uso di soli dati reali incompleti. Infine, Hasan et al. [23] indicano la generazione parametrica come uno dei metodi più efficaci per riprodurre fedelmente la fisiologia umana nei dispositivi wearable.

Sulla base di questi principi, è stata implementata in MATLAB la funzione `genera_sessione` per simulare la traiettoria fisiologica di un soggetto ASD. Ogni sessione ha una durata di 500 minuti (1000 campioni con frequenza di 30 secondi), scelta mirata a riprodurre una finestra di osservazione realistica, paragonabile a una giornata scolastica o a un turno clinico. Questa estensione temporale è fondamentale per stabilire una baseline statistica solida, permettendo al sistema di distinguere le normali variazioni quotidiane dai precursori della crisi. La funzione riproduce l'evoluzione di quattro parametri verso l'evento acuto:

- **EDA e attivazione simpatica:** Per simulare il livello tonico di base, la sudorazione è stata modellata tramite la **funzione di Bates**, che calcola la media di più variabili uniformi per ottenere una distribuzione a campana limitata e controllata tra 2 e 5 μS . Si è evitato il rumore gaussiano poiché presenterebbe code statistiche con valori negativi, fisicamente impossibili per la conduttanza cutanea. Su

questa baseline sono stati sovrapposti 30 picchi fasici spontanei (SCR) modellati come semi-onde sinusoidali per riprodurre la normale risposta galvanica a riposo [24] e, in prossimità della crisi, una rampa di crescita esponenziale ($+8 \mu S$). Questo profilo replica l'iper-attivazione simpatica tipica del *meltdown*, portando la pendenza del segnale (*EDA Slope*) al suo apice nella fase di rischio.

- **HRV e crollo vagale:** La variabilità cardiaca (RMSSD) è stata modellata partendo da un valore medio di 50 *ms*. Nella fase pre-crisi, il segnale subisce un calo lineare del 30%, soglia scelta per distinguere lo stress significativo dalle fluttuazioni quotidiane. Durante l'evento acuto, il valore crolla al 50% rispetto alla baseline, simulando il collasso della capacità di autoregolazione documentato da Hödl et al. [40]. La transizione è stata resa realistica mediante un'interpolazione lineare e l'aggiunta di rumore gaussiano; "sporcare" i dati costringe l'algoritmo a diventare robusto e a riconoscere l'allarme anche in condizioni rumorose tipiche dei sensori reali.
- **ACC e blocco motorio:** L'accelerometria è stata impostata su una baseline di 1.2 *g* (somma della gravità costante e dei micromovimenti di *stimming* del soggetto). Al momento della crisi, il segnale viene ridotto tramite un moltiplicatore di 0.4, subendo un crollo istantaneo. Questo riflette lo *shutdown* motorio descritto in letteratura che spesso anticipa l'esplosione della crisi. Per il modello, una variazione così marcata è l'input ideale per la CPA, permettendo di individuare il momento esatto in cui i parametri deviano dalla baseline. Anche per questo segnale è stato aggiunto un rumore gaussiano calibrato per riflettere l'instabilità dei dati *wearable*.

8.3 Il primo stadio: Change Point Analysis e validazione clinica multimodale

Questo primo stadio dell'architettura ibrida rappresenta il ponte tra l'elaborazione matematica dei segnali e la loro interpretazione clinica, un aspetto che ha richiesto un lavoro di integrazione portato avanti insieme alla collega. In questa prima fase, il si-

stema ha come scopo quello di individuare l'istante temporale in cui la fisiologia del soggetto varia rispetto allo stato di quiete.

Come prima cosa, il sistema analizza il segnale relativo all'attività elettrodermica che è stato inizialmente processato per gestire le continue variazioni fisiologiche, anche a riposo, dovute alla natura stessa del processo di sudorazione; dunque, per evitare che un normale picco elettrodermico fosse scambiato per l'inizio di una crisi, è stata applicata una media mobile con una finestra di 5 campioni in modo tale da smussare le fluttuazioni rapide e lasciare emergere solo la tendenza clinica di fondo. Una volta pulito il segnale, interviene la CPA, che però presenta un limite matematico non trascurabile: cerca sempre la variazione più grande in assoluto; dunque, fornire in input l'intera sessione di 1000 campioni, avrebbe comportato segnalazioni dell'elevato crollo dell'EDA avvenuto subito dopo la crisi, nella fase di recupero. Ma questo sarebbe andato ben oltre l'obiettivo del presente lavoro, che consisteva nel trovare il *lead time* e dunque la sottile rampa di crescita che precede l'evento di sovraccarico. Per superare questo limite si è deciso di vincolare la ricerca del *change point* ai primi 900 campioni, trascurando l'ultimo 10% della sessione. In questo modo la CPA viene forzata a concentrarsi esclusivamente sulla fase di accumulo dello stress.

Non è sufficiente però individuare solo il *change point* sull'EDA, dal momento che, come si evince dalla pratica clinica, un rumore forte o un semplice spavento possono far aumentare l'EDA generando falsi allarmi. Dalla collaborazione con la collega è nata dunque l'idea di effettuare una validazione clinica, ovvero, nel momento esatto individuato dalla CPA, l'algoritmo analizza simultaneamente i dati provenienti da tutti gli altri sensori, applicando quattro regole calibrate sulla fisiologia individuale del soggetto e nata dalla precedente analisi effettuata sui dati reali.

- **Regola 1 – EDA alta:** La sudorazione tonica deve superare l'85° percentile della sua distribuzione giornaliera. Questo è definito come il segno primario di un sovraccarico acuto.
- **Regola 2 – Slope EDA alta:** La velocità con cui l'EDA sta salendo deve superare il 75° percentile. Questo indica che la tensione non è solo alta, ma sta crescendo rapidamente.

- **Regola 3 – Crollo motorio:** L’accelerometria deve scendere sotto l’80% del suo valore mediano. Questo parametro traduce il blocco motorio e la brusca interruzione dei comportamenti di autoregolazione (*stimming*).
- **Regola 4 – Crollo HRV:** La variabilità cardiaca deve scendere sotto il 90% della sua mediana, segno inequivocabile che il sistema nervoso simpatico ha preso il sopravvento, causando il ritiro del controllo vagale parasimpatico.

È importante sottolineare che per le soglie di ACC e HRV si è scelto di utilizzare la mediana anziché la classica media matematica. Questo perché basta un singolo movimento brusco del braccio per registrare un picco anomalo che falserebbe del tutto la media matematica. La mediana, al contrario, ignora naturalmente questi sbalzi isolati, garantendo una soglia di allarme molto più stabile e veritiera.

Il superamento di questo primo stadio è governato da una logica booleana “AND + OR”, pensata per imitare il processo di diagnosi differenziale di un medico. L’anomalia individuata viene validata come una sospetta pre-crisi solo se la regola 1 è verificata e se contemporaneamente è verificata almeno un’altra tra le regole 2, 3 o 4. In questo modello è dunque la concordanza di più biomarcatori indipendenti a certificare il reale inizio del sovraccarico sensoriale.

8.4 Il secondo stadio: predizione tramite Random Forest

Se il primo stadio dell’architettura ha come scopo quello di individuare il punto di rottura fisiologico, il secondo stadio ha come obiettivo quello di trasformare questa anomalia in una predizione clinica quantificabile. Dall’analisi precedente sui dati reali si è dunque deciso di implementare un classificatore *ensemble* Random Forest, addestrato per calcolare in tempo reale la probabilità che il soggetto vada incontro a un sovraccarico emotivo, ancor prima che questo sfoci in una crisi comportamentale visibile.

Una delle criticità maggiori affrontate nell’addestramento di questo modello predittivo è stata l’assenza di eventi critici nel dataset reale disponibile. Generalmente,

il problema dello sbilanciamento delle classi vien gestito a livello algoritmico tramite l'applicazione di una matrice dei costi o metodi di ricampionamento (come lo SMOTE). Tuttavia, questi approcci classici richiedono la presenza di almeno un numero minimo di campioni della classe minoritaria da cui l'algoritmo possa apprendere l'anomalia.

In una condizione clinica come quella affrontata in cui si hanno zero crisi registrate, l'algoritmo è del tutto privo di una *ground truth* di riferimento e, di conseguenza, matematicamente non si può penalizzare un errore di classificazione, o interpolare dati artificiali partendo dal nulla, se il modello non ha mai visto la fisiologia di un determinato evento. Tale ostacolo strutturale ha precluso l'impiego dei tradizionali metodi di bilanciamento, rendendo lo sviluppo dell'architettura ibrida un passaggio metodologico obbligato.

Ecco perché i dati sintetici (descritti nel paragrafo 8.2) non sono stati un semplice espediente per migliorare le prestazioni, ma l'unica vera soluzione al problema. Sebbene l'analisi sui soli dati reali avesse già permesso di predire il sovraccarico emotivo, visto che i dati reali non offrivano esempi di crisi conclamate da far studiare all'algoritmo, sono stati costruiti artificialmente tramite la *data augmentation*. Nel dataset di addestramento sono state inserite le simulazioni cliniche che mancavano così che il Random Forest si trovasse davanti a un set di dati equilibrato, grazie al quale ha potuto studiare sia le ore di calma sia i momenti critici, imparando a riconoscere quei *pattern* fisiologici di allarme che, altrimenti, non avrebbe mai potuto vedere.

Per la fase di *training* vera e propria, è stato utilizzato un approccio di tipo *bagging* (termine che nasce dalla contrazione di *bootstrap aggregating*). Questa tecnica di apprendimento *ensemble* consiste nell'estrarre casualmente diversi sottoinsiemi dal dataset di partenza per addestrare modelli multipli e paralleli. Nello specifico, l'algoritmo è stato strutturato per generare una "foresta" di 50 alberi decisionali indipendenti. In questo modo, la predizione non viene effettuata da un unico e rigido modello matematico, ma il sistema consulta 50 "alberi" e ne aggrega i risultati tramite un voto di maggioranza. Questa tecnica permette di ridurre l'errore e aumentare la stabilità e l'affidabilità della previsione finale. Inoltre, uno dei vantaggi dell'impiego di un Random Forest risiede nella sua natura intrinsecamente non lineare. Infatti, a differenza dei classificatori statistici classici, che cercano di tracciare linee di separazione nette tra i

dati, gli alberi decisionali ragionano attraverso soglie a cascata. Questa struttura logica permette di tenere conto delle complesse interazioni fisiologiche che caratterizzano una crisi; l'algoritmo, cioè, non guarda i singoli segnali in modo isolato, ma impara ad avere una visione d'insieme: un semplice innalzamento della sudorazione (EDA) non basta per far scattare un vero allarme, ma lo diventa solo se, in quello stesso identico momento, il soggetto si blocca a livello motorio (ACC) e il suo sistema nervoso va in tilt perdendo il controllo vagale (HRV).

Per garantire che le metriche di accuratezza del modello non fossero penalizzate dal fenomeno di overfitting, la validazione è stata eseguita tramite k-fold cross-validation impostando $k = 5$. Quindi, il dataset è stato diviso in 5 sottoinsiemi: il modello è stato addestrato iterativamente su quattro di essi (l'80% dei dati) e testato sull'ultimo (il 20%). Impostando $k = 5$ ci si è assicurati che, in ogni ciclo, il *test set* fosse abbastanza grande (il 20%, appunto). Se il dataset fosse stato diviso in troppe parti (ad esempio $k = 10$), la porzione di test sarebbe stata troppo piccola e le metriche di valutazione avrebbero rischiato di essere falsate a causa di un singolo artefatto o di un movimento brusco isolato. Con questo, le percentuali di sensibilità e specificità ottenute riflettono in modo affidabile la reale capacità dell'algoritmo di generalizzare su dati fisiologici mai osservati prima.

Nella fase di addestramento dell'algoritmo, si è dovuto definire cosa il modello dovesse imparare. Se si fosse etichettato come "rischio" solo l'istante esatto dell'esplosione comportamentale, il modello si sarebbe limitato a confermare un evento già in corso. L'obiettivo primario di questo studio è invece garantire un tempo di anticipo sufficiente per permettere di intervenire sul sovraccarico emotivo in fase di accumulo e tempestivamente. Nel codice MATLAB, questo concetto è stato tradotto attraverso una tecnica di *time-shifting*. Una volta individuato il momento esatto in cui avviene la crisi, la tecnica citata è stata applicata traslando la finestra di allarme all'indietro nel tempo. Poiché l'algoritmo analizza i dati dei sensori a blocchi, ovvero mediante finestre mobili di 30 secondi ciascuno, è stato imposto di fare un salto all'indietro di 3 *step*, posizionandosi esattamente un minuto e mezzo prima dell'evento. In termini pratici, l'algoritmo è stato istruito per analizzare le variazioni fisiologiche antecedenti l'evento acuto, assegnando alla classe positiva la finestra temporale dei 90 secondi

precedenti la crisi. Avviene dunque uno spostamento del target predittivo, al fine di garantire non l'identificazione del *meltdown*, ma l'intercettazione dell'accumulo dello stress fisiologico, rilevando il sovraccarico emotivo interno che potrebbe sfociare in una crisi aggressiva.

Infine, si è prestata attenzione alla trasparenza del modello, ovvero capire quali metriche guidassero le sue decisioni, proseguendo quindi con l'estrazione della *feature importance* calcolando matematicamente l'indice di Gini, per osservare quale parametro avesse maggior peso. L'algoritmo, in risposta, restituisce una vera e propria classifica, in base all'influenza che ogni sensore ha avuto nella decisione. Questo consente di comprendere se l'allarme è scattato ad esempio a causa di un blocco motorio improvviso o per un forte aumento dell'EDA, e dunque quale reazione fisiologica sta portando il soggetto verso il sovraccarico così da poter calibrare l'intervento.

8.5 Valutazione delle performance e analisi dei risultati

Al fine di validare l'efficacia dell'architettura proposta, il modello è stato sottoposto a una rigorosa fase di *test*. Come illustrato nelle sezioni precedenti, la fase di *training* è avvenuta sui dati generati per simulare un soggetto noto. Per verificare la reale capacità predittiva del sistema, le sue performance sono state successivamente misurate su una sessione di 500 minuti appartenente a un soggetto "ignoto" (fase di *test*), ovvero su dati fisiologici che il Random Forest non aveva mai analizzato prima. I risultati ottenuti confermano la solidità del sistema ibrido su molteplici fronti.

8.5.1 L'identificazione del rischio e il *lead time*

L'analisi visiva della sessione di *test* dimostra il funzionamento dell'architettura a cascata: il primo stadio non supervisionato ha individuato correttamente il *change point* in corrispondenza dell'inizio della rampa di crescita esponenziale dell'EDA. Successivamente, l'algoritmo ha confermato l'anomalia rilevando il crollo sincrono dell'accelerometria e dell'HRV. A questo punto, il Random Forest ha iniziato a elaborare la

probabilità continua di crisi. L'anticipo con cui il sistema lancia l'allarme rispetto al momento effettivo in cui la crisi esplode visibilmente, definito dal *lead time*, è stato un risultato significativo: il modello ha superato la soglia di rischio del 25% con un anticipo di 2.5 minuti. Questo dato si inserisce perfettamente nella finestra temporale ideale (1-3 minuti) indicata dalla letteratura come il tempo clinicamente utile e necessario a un educatore per attuare strategie calmanti e prevenire il *meltdown*.

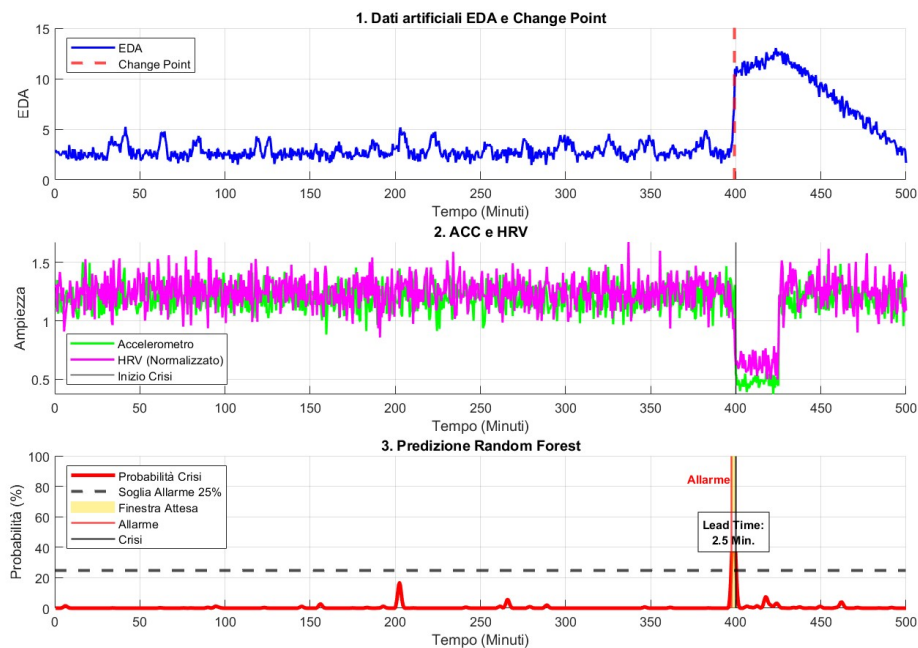


Figura 8.2: Visualizzazione macro della sessione di test di 500 minuti. Pannello 1: segnale EDA (blu) con Change Point rilevato (linea rossa tratteggiata). Pannello 2: segnali ACC (verde) e HRV (magenta) con crollo sincrono all'inizio della crisi (linea nera verticale). Pannello 3: probabilità di crisi prevista (linea rossa) rispetto alla soglia del 25% (linea orizzontale tratteggiata) e finestra di attesa (area gialla), con allarme preventivo e calcolo del Lead Time.

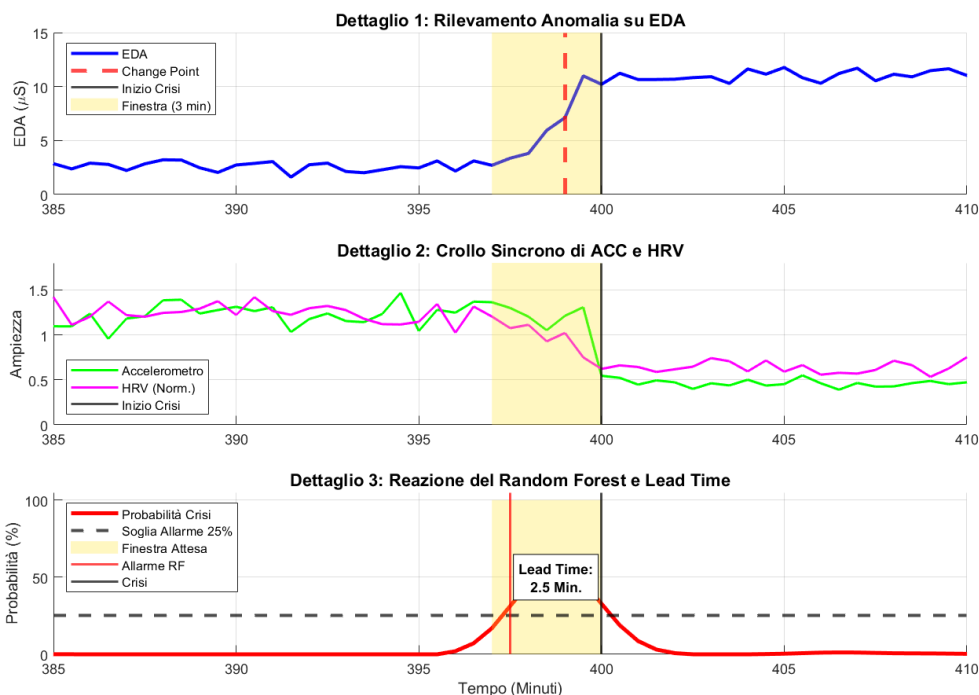


Figura 8.3: Dettaglio micro della finestra temporale critica (385-410 minuti). Lo zoom permette di analizzare la dinamica dettagliata degli eventi. Pannello 1: rampa di crescita dell'EDA con Change Point e finestra pre-crisi evidenziata. Pannello 2: crollo sincrono di ACC e HRV all'inizio della crisi. Pannello 3: impennata della probabilità prevista, attraversamento della soglia e attivazione dell'allarme preventivo, con Lead Time di 2.5 minuti.

8.5.2 La Matrice di Confusione

L'affidabilità statistica del modello è stata misurata valutando la capacità di classificazione su tutti i 1000 campioni della sessione di *test*. In questo scenario di monitoraggio (circa 8 ore totali), l'algoritmo ha dovuto distinguere tra lo stato di quiete e i 7 campioni critici corrispondenti alla finestra di circa 3 minuti, che precedono l'evento acuto.

Dall'analisi della matrice di confusione è emerso che su 993 momenti appartenenti alla fase di calma, il sistema non ha mai generato segnalazioni errate (Veri Negativi = 993, Falsi Positivi = 0); e che dei 7 momenti che compongono il sovraccarico verso la crisi, l'algoritmo ne ha identificati correttamente 6 (Veri Positivi = 6). È sfuggito un solo istante di transizione (Falso Negativo = 1), situato proprio all'inizio della rampa, quando i segnali fisiologici erano matematicamente ancora troppo simili alla *baseline* per far scattare la soglia di allerta.

In termini di metriche globali, il sistema ha raggiunto una Sensibilità dell'85,7% e una Specificità e Precisione del 100%. È doveroso precisare che il raggiungimento

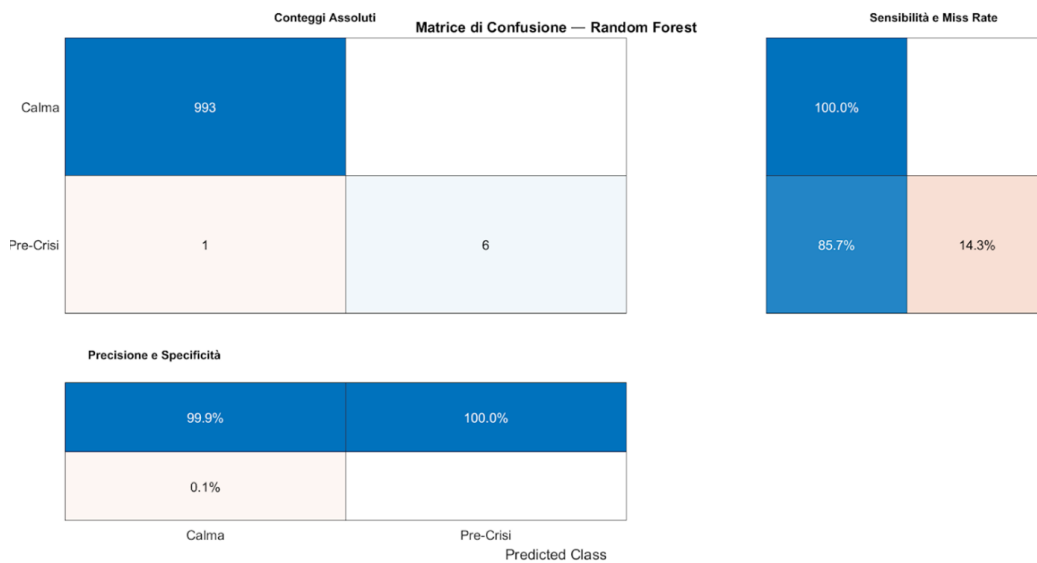


Figura 8.4: Matrice di Confusione del modello Random Forest calcolata sulla sessione di *test*. I pannelli illustrano i conteggi assoluti (a sinistra) e le metriche percentuali derivate.

di una precisione impeccabile in questa fase di *test* è la conseguenza diretta dell'uso di dati sintetici, che fornisce un ambiente di validazione più controllato rispetto alla variabilità caotica dei segnali reali, e dall'efficacia dell'architettura a due stadi estremamente severa. Infatti, la logica di validazione incrociata tra i tre sensori (EDA, HRV e ACC) funge da filtro, impedendo a semplici artefatti o rumori isolati di attivare l'allarme. Questa scelta privilegia la stabilità rispetto alla massima sensibilità, con l'obiettivo prioritario di proteggere il *caregiver* dal cosiddetto stress da falsi allarmi. Garantire che ogni segnalazione corrisponda a un rischio reale è ciò che rende il modello un supporto concreto, affidabile e realmente applicabile nella pratica clinica quotidiana.

8.5.3 Capacità discriminativa globale: curva ROC e AUC

Per valutare la validità del classificatore indipendentemente dalla singola soglia di allarme scelta (25%), è stata tracciata la curva ROC (*Receiver Operating Characteristic*). Questo strumento analizza il bilanciamento tra la sensibilità e il tasso di falsi positivi su ogni possibile scenario operativo, fornendo una panoramica completa riguardante la robustezza del modello.

Il valore di sintesi, l'AUC (*Area Under the Curve*), ha raggiunto un punteggio di 0.85. Tale valore indica che il modello ha l'85% di probabilità di assegnare corretta-

mente un punteggio di rischio superiore a un campione di precrisi rispetto a uno di calma. In letteratura, valori di AUC superiori a 0.80 sono considerati indice di un potere predittivo molto buono; inoltre, tale valore conferma la capacità dell’algoritmo di isolare l’attivazione simpatica anomala dal normale rumore di fondo dei segnali biologici. Un altro aspetto rilevante è che la curva si posiziona nettamente al di sopra della diagonale di riferimento ($AUC = 0.50$), che rappresenta un classificatore casuale e ciò prova che le predizioni derivano da un reale apprendimento dei *pattern* fisiologici e non da semplici fluttuazioni casuali dei dati.

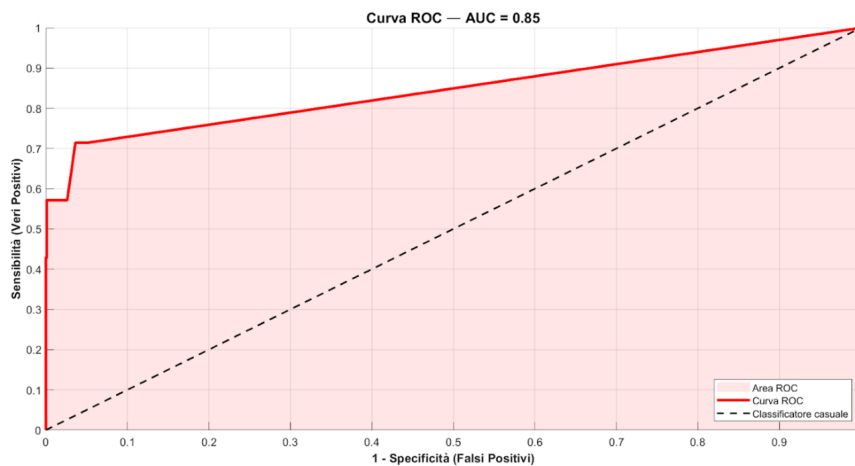


Figura 8.5: Curva ROC del classificatore Random Forest. L’area sottesa alla curva (AUC) è pari a 0.85, dimostrando un’ottima capacità discriminativa del modello posizionandosi nettamente al di sopra della diagonale del classificatore casuale ($AUC = 0.50$).

8.5.4 L’Importanza delle feature

Da questa analisi, è emerso che la variabile determinante è la pendenza dell’EDA, il cui peso è circa dieci volte superiore rispetto agli altri parametri. L’algoritmo ha dunque appreso che il marcatore più affidabile della crisi non è il livello assoluto di sudorazione, che può variare naturalmente per fattori ambientali o fisici, ma la rapidità della sua ascesa. Matematicamente, l’indice di Gini premia la capacità della pendenza di discriminare nettamente tra lo stato di quiete e l’inizio del sovraccarico. In sostanza, il modello non si limita a osservare quanto il soggetto sia attivato, ma monitora la velocità di accumulo dello stress, identificando l’imminenza della crisi proprio nel momento

in cui alla crescita dell'EDA si associa il simultaneo ritiro del sistema parasimpatico (HRV), segno del collasso dei meccanismi di autoregolazione.

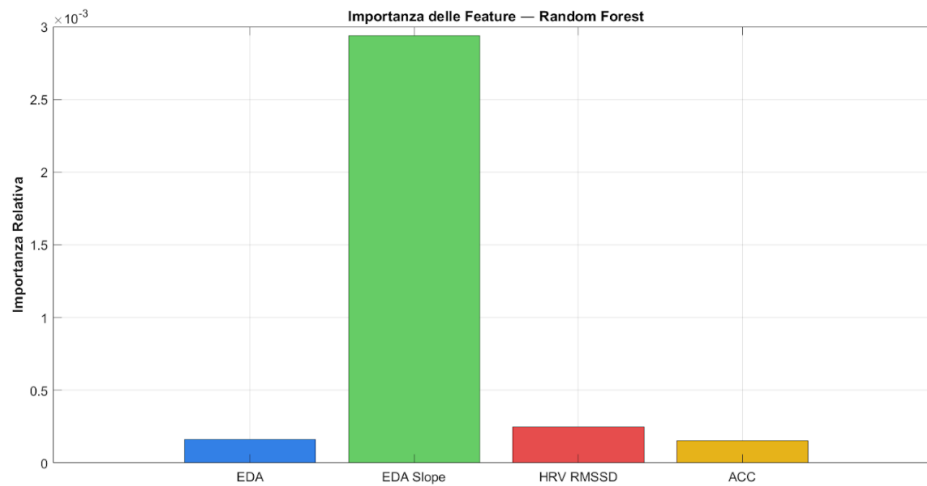


Figura 8.6: Importanza relativa delle *feature*. Il grafico evidenzia in modo netto come la pendenza dell'EDA sia la variabile predominante per la classificazione del rischio, superando di gran lunga il peso del valore assoluto dell'EDA, dell'HRV e dell'accelerometria (ACC).

8.5.5 Capacità di generalizzazione

L'ultimo passaggio della validazione ha riguardato la verifica della tenuta del modello su dati nuovi, per escludere il rischio di *overfitting*. Per farlo, sono state messe a confronto le prestazioni del Random Forest in due momenti distinti:

- **Fase di Training (soggetto noto):** Il modello doveva riconoscere la crisi nello stesso soggetto su cui si era addestrato. In questo caso, la probabilità di rischio ha raggiunto un picco del 68%.
- **Fase di Test (soggetto ignoto):** Sono stati sottoposti al modello i dati di un secondo soggetto simulato, le cui variazioni fisiologiche erano matematicamente diverse dalle prime. In questa prova "cieca", il picco di probabilità si è attestato al 58%.

Il fatto che i risultati calino solo del 10% passando da un soggetto noto a uno sconosciuto è un segno di solidità; inoltre, è abbastanza naturale questo calo dal momento che, appunto, il modello affronta un dato mai visto prima. Tuttavia, restando

al 58% (ben oltre la soglia d'allarme del 25%), dimostra di aver acquisito capacità di ragionamento.

Questo test dimostra che l'algoritmo non ha semplicemente memorizzato il *dataset* (evitando il fenomeno dell'*overfitting*), ma ha generalizzato la regola biologica sottostante: a prescindere dal soggetto in *input*, il rischio clinico è reale quando un'elevata attivazione sudomotoria si scontra con un crollo della variabilità cardiaca. Questa flessibilità diagnostica deriva direttamente dall'implementazione delle soglie relative a percentili. Svincolandosi dalla ricerca di valori assoluti rigidi e predefiniti, il modello calibra matematicamente la propria "normalità" sulla *baseline* del singolo individuo. Rilevando l'allarme solo di fronte a deviazioni critiche dall'equilibrio unico del soggetto, l'infrastruttura algoritmica si eleva a strumento di protezione su misura, capace di dare voce al linguaggio interno e silenzioso di ogni corpo.

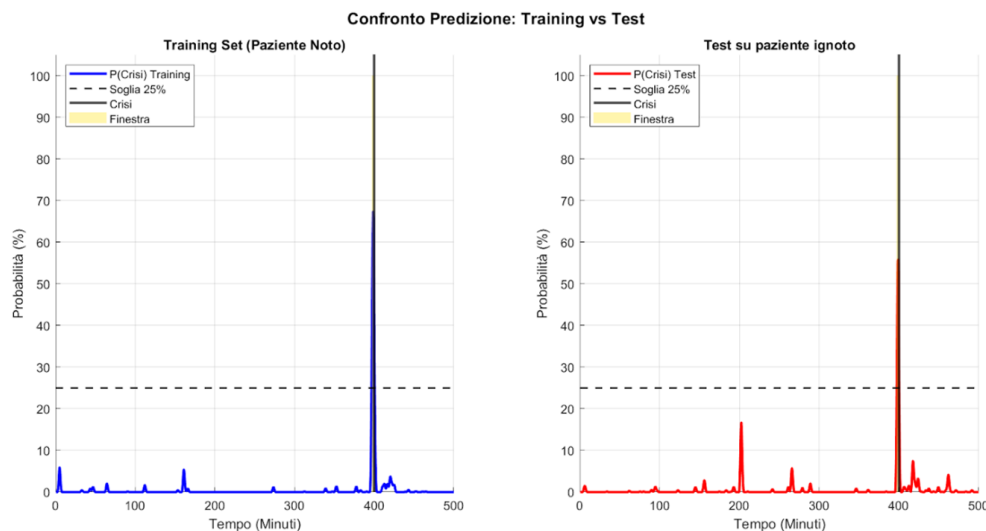


Figura 8.7: Confronto delle prestazioni predittive del modello tra la fase di *Training* (soggetto noto, a sinistra) e la fase di *Test* (soggetto ignoto, a destra).

8.6 Analisi di robustezza al rumore (+6 dB)

Nei dispositivi indossabili, il segnale registrato è intrinsecamente soggetto a interferenze che lo degradano. Tali disturbi derivano principalmente da artefatti da movimento e da micro-fluttuazioni fisiologiche legate a processi come la termoregolazione, che non hanno alcun significato clinico rispetto allo stress o alle crisi. Nel codice, questa interferenza non è stata simulata in modo generico, ma generando specifiche componenti

stocastiche per ogni segnale. È stato quindi applicato un raddoppio dell'ampiezza di tali componenti che, matematicamente, corrisponde a un aumento di circa 6 decibel, calcolato tramite la formula standard:

$$\Delta dB = 20 \cdot \log_{10} \left(\frac{A_{nuovo}}{A_{iniziale}} \right) \quad (8.1)$$

$$\Delta dB = 20 \cdot \log_{10}(2) \approx 6.02 \text{ dB} \quad (8.2)$$

Applicando questo moltiplicatore, i parametri di rumore del nuovo soggetto virtuale sono stati così ricalibrati:

- **EDA:** Il rumore basale è stato generato utilizzando una distribuzione di Bates (calcolata come media di cinque distribuzioni uniformi indipendenti), che garantisce fluttuazioni realistiche e limitate. L'ampiezza massima di tale rumore è stata portata a $6 \mu S$ (rispetto ai $3 \mu S$ iniziali);
- **HRV:** Il disturbo è stato generato con una distribuzione gaussiana, la cui deviazione standard è stata portata a 10 ms (rispetto ai 5 ms iniziali);
- **ACC:** Anche in questo caso è stato aggiunto un rumore gaussiano, raddoppiando la deviazione standard a 0.2 g (rispetto agli 0.1 g originali).

Per comprendere l'impatto di questa operazione, è fondamentale richiamare la definizione del rapporto segnale/rumore (SNR), espresso in decibel secondo la formula:

$$SNR_{dB} = 10 \cdot \log_{10} \left(\frac{\text{Potenza Segnale}}{\text{Potenza Rumore}} \right) \quad (8.3)$$

Nel modello di partenza, il rapporto segnale/rumore (SNR) si attestava su valori ottimali. Utilizzando la formula logaritmica applicata ai dati ($SNR = 20 \cdot \log_{10}(\mu/\sigma)$), il modello iniziale registrava un SNR di circa +20.0 dB per l'HRV e di +21.6 dB per l'accelerometro (ACC). Questo garantiva un segnale "pulito" in cui i pattern clinici emergevano chiaramente. Poiché raddoppiando l'ampiezza del rumore la sua potenza quadruplica, questa operazione si traduce in una riduzione dell'SNR di esattamente 6

dB, causando dunque un abbassamento strutturale della proporzione tra il segnale utile (la rampa pre-crisi) e il disturbo di fondo.

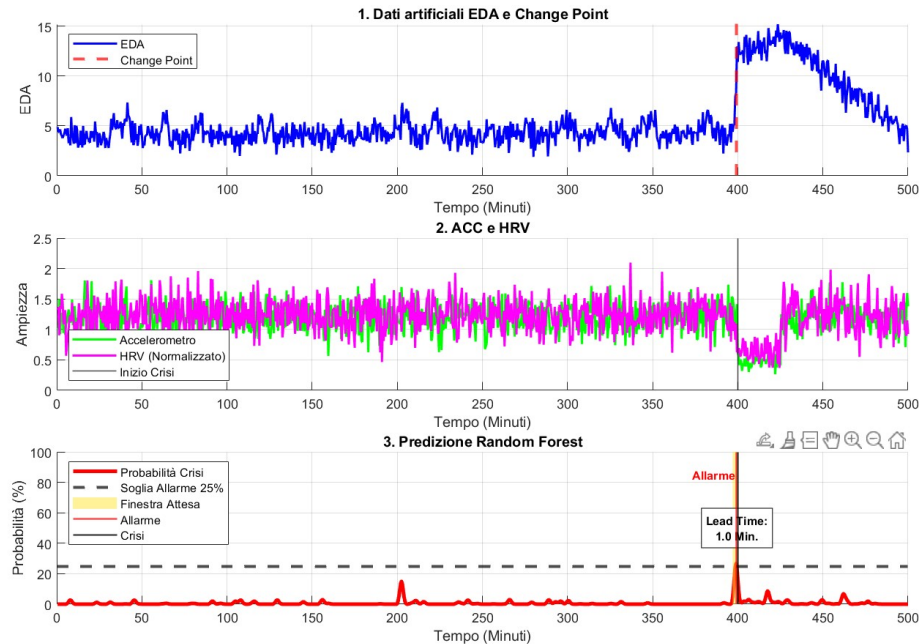


Figura 8.8: Dati artificiali e Change Point (in alto), segnali ACC e HRV (al centro) e predizione Random Forest (in basso) in condizioni di rumore maggiorato (+6 dB). È evidente il crollo predittivo, con l'allarme ritardato e un Lead Time ridotto a 1.0 minuto.

Sottoponendo questi dati rumorosi all'architettura ibrida, dall'analisi visiva sono stati osservati da subito segnali caotici, con la linea di base dell'EDA costituita da picchi casuali fino a $8 \mu S$ simulanti artefatti dovuti a sfregamento del sensore. Di fronte a questo nuovo scenario, i due stadi dell'architettura ibrida hanno mostrato due tipi di risposte diverse:

1. **Change Point Analysis:** Ha continuato a posizionare la linea di rottura esattamente al minuto 400. Questo risultato conferma che la scelta di applicare una finestra di media mobile a monte del rilevamento ha attenuato le oscillazioni aggiuntive, permettendo all' algoritmo non supervisionato di individuare la rottura strutturale senza alcun ritardo.
2. **Random Forest:** Il classificatore supervisionato, al contrario, ha perso la sua capacità di anticipazione predittiva. La curva di probabilità, originariamente

piatta, è diventata instabile, generando fluttuazioni e falsi picchi fino al 20% durante le fasi di calma. Sebbene il modello non abbia generato falsi allarmi, la sua capacità di riconoscere il pericolo in tempo utile è svanita; dall'analisi della sensibilità e della matrice di confusione si è osservato che i veri positivi sono scesi da 6 a 1, mentre i falsi negativi (campioni di precrisi scambiati per calma) sono saliti a 6. La sensibilità è conseguentemente crollata dall'85,7% al 14,3%.

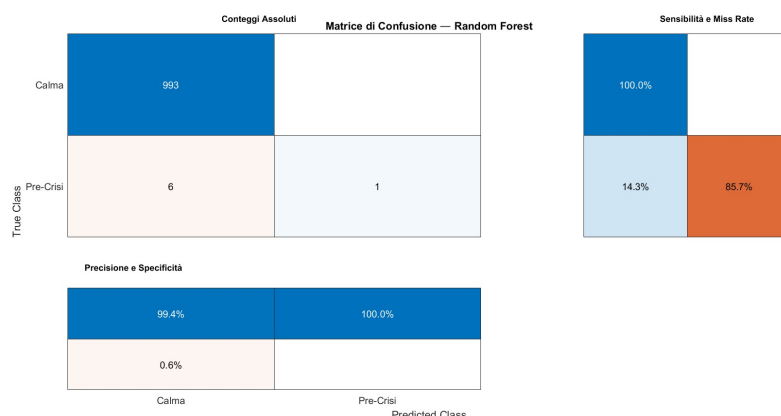


Figura 8.9: Matrice di confusione del Random Forest in condizioni di rumore maggiorato (+6 dB). Si osserva il drastico crollo della sensibilità al 14,3%, con 6 campioni di pre-crisi su 7 non riconosciuti (falsi negativi).

Per quanto riguarda la *lead time*, esso è sceso da 2.5 minuti a un solo minuto, per una perdita di 1.5 minuti di finestra di intervento (figura 8.8). Anche la capacità discriminante globale ne ha risentito, con una AUC diminuita da 0.85 a 0.76 (figura 8.10).

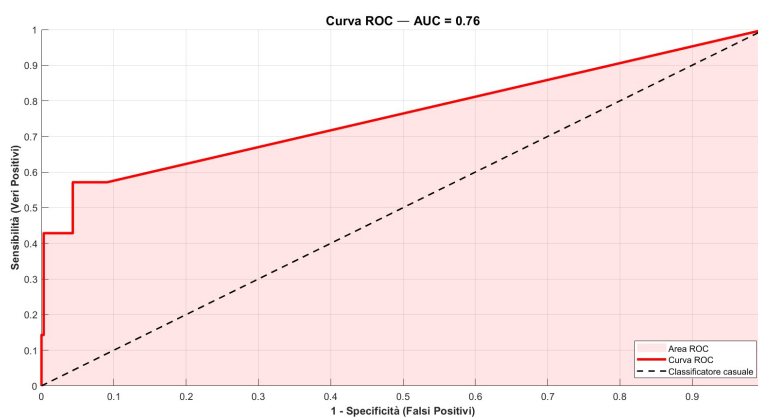


Figura 8.10: Curva ROC del modello Random Forest sottoposto ai dati con rumore maggiorato. L'Area Sottesa alla Curva (AUC) subisce una flessione, scendendo al valore di 0.76, a conferma della perdita di capacità discriminante causata dalle fluttuazioni ad alta frequenza.

Dall'estrazione della feature importance è stato possibile osservare la causa matematica di questo abbassamento drastico delle performance: l'importanza della pendenza dell'EDA, che nel modello originale era il predittore dominante, si è ridotta di oltre il 60%. Matematicamente, calcolare una pendenza equivale a derivare il segnale e, la derivazione, per sua natura, amplifica enormemente le fluttuazioni ad alta frequenza. Di conseguenza, con l'aggiunta di rumore a +6 dB, l'*EDA Slope* è diventata caotica, celando l'informazione utile del segnale; di conseguenza, l'algoritmo predittivo, privato della sua variabile più importante, non è più riuscito a distinguere la rampa pre-crisi dalla rumorosa linea di fondo, attivando l'allarme solo quando i segnali grezzi sono crollati in modo evidente a ridosso dell'evento.

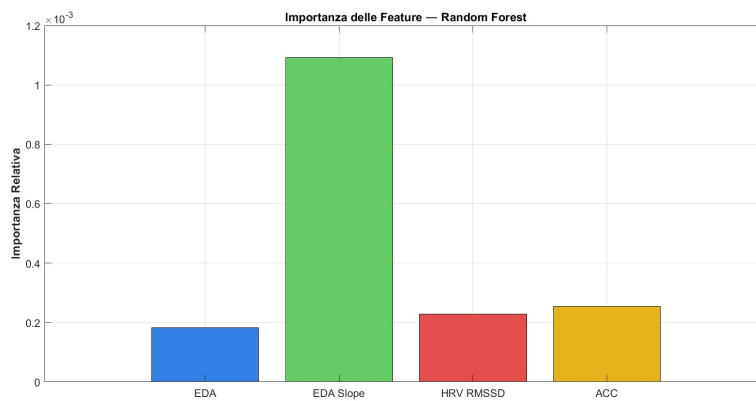


Figura 8.11: Istogramma della feature importance per il modello Random Forest addestrato sui dati rumorosi.

Infine, questo crollo di capacità predittiva è stato confermato dall'analisi di generalizzazione (figura 8.12). Mentre sui dati noti di addestramento il modello riesce comunque ad individuare la crisi mantenendo un picco superiore al 50%, sui dati di verifica (soggetto ignoto) la curva di probabilità risulta estremamente instabile e fatica a superare la soglia operativa del 25%. Questo dimostra che il rumore ad alta frequenza, mascherando l'andamento dell'EDA, inficia drasticamente sulla capacità dell'algoritmo di generalizzare i pattern fisiologici su dati mai visti.

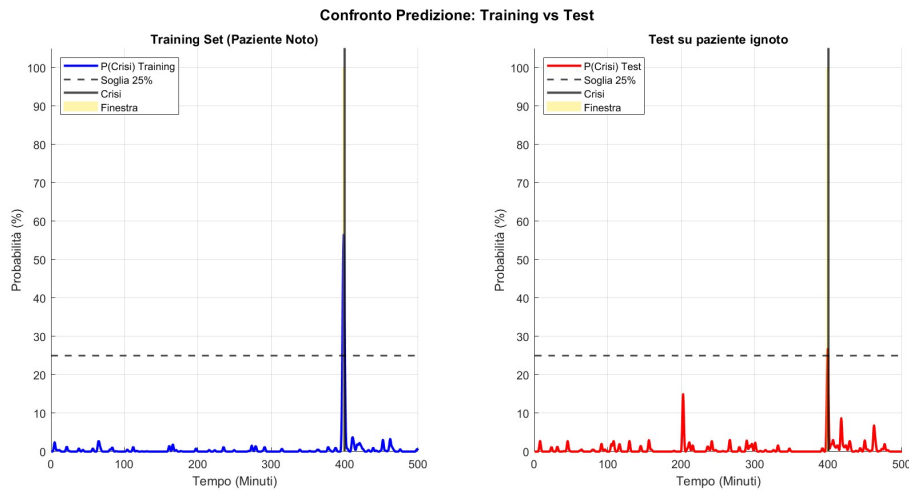


Figura 8.12: Confronto della probabilità predittiva tra i dati di addestramento (Training Set, a sinistra) e i dati di verifica (Test Set, a destra) in presenza di rumore maggiorato.

8.6.1 Ottimizzazione del segnale tramite pre-processing

Come evidenziato dall'analisi precedente, il crollo della sensibilità risiedeva interamente nella scarsa qualità della *feature* in quel momento più rilevante per l'algoritmo: la pendenza dell'EDA. Aumentando il rumore di +6 dB, il segnale oscilla rapidamente e caoticamente da un campione all'altro, e, calcolare matematicamente la derivata su un segnale così disturbato (tramite la funzione *gradient* di MATLAB), vuol dire calcolare la differenza tra una oscillazione casuale e quella successiva amplificando i piccoli disturbi casuali, trasformandoli in enormi picchi che confondono l'algoritmo rispetto all'informazione utile. Per risolvere questa problematica e ripristinare le capacità predittive del modello, sono state effettuate due modifiche mirate alla funzione di generazione e processamento dei dati, con l'obiettivo di agire sulla qualità dei segnali in ingresso.

La prima modifica riguarda l'utilizzo di un filtro a media mobile per rendere la pendenza nuovamente leggibile. Il segnale è stato dunque filtrato prima di calcolarne la derivata: è stata introdotta una media mobile a 5 campioni (tramite la funzione *movmean*). Questa operazione sostituisce ogni campione con la media dei 5 campioni adiacenti (che corrispondono a una finestra temporale di 2.5 minuti). In parole pratiche, invece di considerare ogni singolo valore del segnale così come viene registrato (che potrebbe essere rumore) il sistema sostituisce quel valore con la media calcolata

tra il valore stesso e i 4 campioni vicini. La scelta di una finestra di 5 campioni deriva dal giusto compromesso tra la lunghezza necessaria a eliminare i disturbi rapidi e casuali, e quella sufficientemente corta da non cancellare la rampa di accumulo dello stress, che è più lenta e graduale. Questo stesso tipo di “pulizia” con 5 campioni viene usata anche nel primo stadio del modello; e utilizzare lo stesso parametro per tutto il sistema garantisce coerenza strutturale e una analisi del segnale con lo stesso grado di precisione. In sostanza, questo filtro agisce come un “filtro passa-basso”. Significa che lascia passare i segnali lenti e importanti (le “basse frequenze”, come lo stress che sale) e blocca invece i segnali rapidi e disordinati (le “alte frequenze”, ovvero il rumore).

Oltre al filtro, è stata apportata una seconda modifica per rendere la simulazione ancora più realistica. Nel modello ideale, l’ampiezza della rampa esponenziale dell’EDA cresceva di un valore fisso (+8 μS). Nel nuovo codice, l’algoritmo è stato modificato per generare un incremento variabile e casuale compreso tra 4 e 10 μS . Nella realtà, infatti, l’intensità del sovraccarico fisiologico varia abbastanza tra episodi diversi e tra soggetti diversi. Questa modifica non è stata introdotta per migliorare direttamente la sensibilità del modello, ma per rendere il contesto di lavoro più realistico e complesso ed evitare il rischio di *overfitting*.

Per testare la validità dell’applicazione di questo pre-processing, sono state analizzate le metriche estratte con i seguenti risultati:

- **Sensibilità:** è tornata all’85,7%. Ripristinata la leggibilità della pendenza, il modello ha ricominciato a classificare correttamente 6 dei 7 campioni della finestra di pre-crisi.
- **Specificità:** si è mantenuta su livelli di eccellenza (99,9%), registrando un solo falso allarme su 993 campioni appartenenti alla fase di calma.

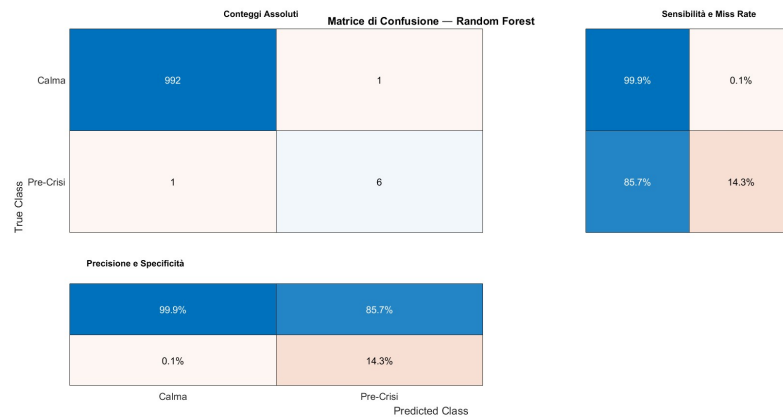


Figura 8.13: Matrice di confusione del modello Random Forest dopo l'applicazione del pre-processing (filtro a media mobile e rampa variabile). Si osserva il ripristino della sensibilità all'85,7% per la classe "Pre-Crisi" e un'alta specificità del 99,9% per la classe "Calma".

- Lead Time:** l'anticipo predittivo è tornato al valore di 2.5 minuti. La curva di probabilità è rimasta stabile e aderente allo zero per l'intera fase basale (con picchi isolati mai superiori al 15%), superando la soglia di allarme del 25% esattamente al minuto 397.5, garantendo l'anticipo necessario per intervenire prima del crollo al minuto 400.

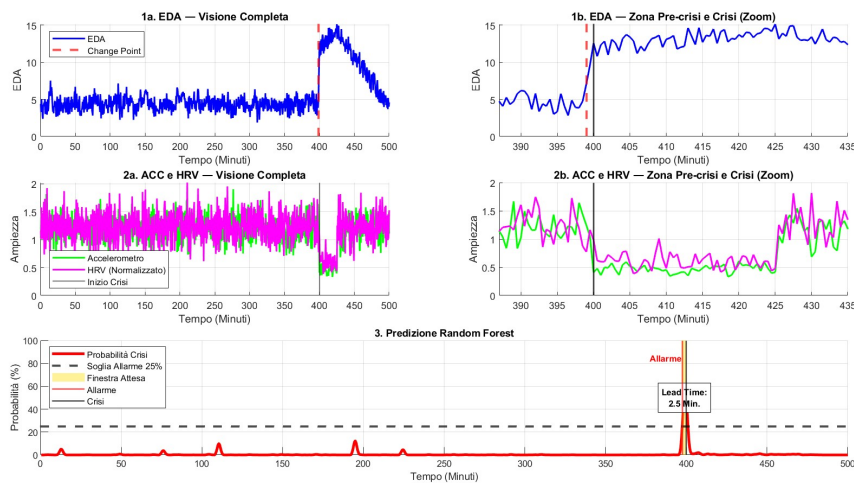


Figura 8.14: Visualizzazione dei segnali fisiologici e della predizione del modello dopo l'applicazione del pre-processing. Nei grafici superiori è visibile l'EDA (con relativo zoom sulla rampa pre-crisi); al centro l'andamento di ACC e HRV. Nel grafico inferiore, la curva di probabilità del Random Forest e il *lead time*

- Capacità discriminante globale:** L'Area sottesa alla curva ROC ha raggiunto il valore di 0.91. Osservando il grafico di questa curva, si nota che la linea non

è morbida e arrotondata, ma sale a “gradini”. Questo accade matematicamente quando l’evento che si sta cercando è un evento raro: avendo a disposizione solo 7 campioni di precrisi su un totale di 1000 campioni di calma, ogni volta che il modello riconosce correttamente uno di quei 7 momenti critici, la linea del grafico fa un balzo netto verso l’alto. Il fatto che il primo gradino sia così alto e ripido dimostra che l’algoritmo ha un’alta confidenza nell’identificare il rischio fin dai primissimi istanti. Inoltre, è importante fare una considerazione oggettiva su questo punteggio così elevato. Il raggiungimento di un’AUC pari a 0.91 è in parte dovuto all’artificialità dei dati che sono stati utilizzati; nonostante sia stato aggiunto rumore, le traiettorie fisiologiche matematiche mantengono sempre una certa regolarità di fondo. Mentre nel mondo reale, i dati sarebbero più caotici. Per questo motivo, se si usasse tale modello su soggetti reali, la forma e i tempi con cui l’EDA aumenta verso la crisi varierebbero in modo molto più imprevedibile da soggetto in soggetto e, logicamente, ci si attenderebbe un’AUC più bassa. Anche la letteratura scientifica di riferimento dimostra chiaramente che, analizzando dati reali per la previsione dei *meltdown*, i modelli migliori e considerati “ottimali” per l’uso clinico raggiungono un’AUC compresa tra 0.83 e 0.87. Il nostro modello, avendo raggiunto oggi lo 0.91, dimostra di avere un ampio margine di sicurezza, presentando una certa robustezza per poter gestire la naturale imprevedibilità del corpo umano e rimanere comunque nella fascia di eccellenza richiesta per l’applicazione clinica.

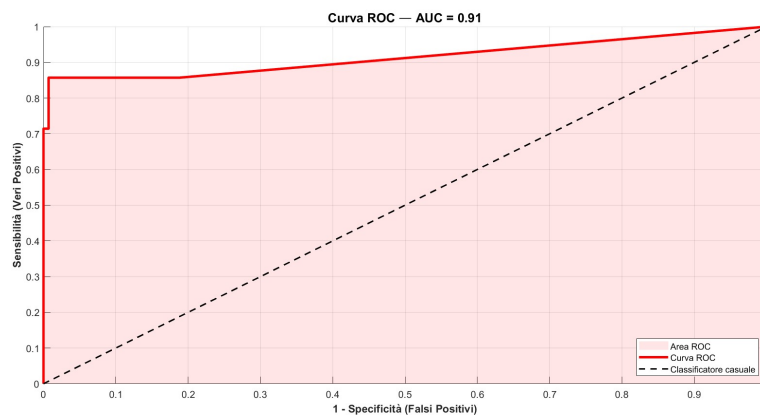


Figura 8.15: Curva ROC del modello ottimizzato.

8.7 Analisi di robustezza temporale: variabilità dell'onset

Finora, le prestazioni del modello sviluppato sono state valutate su un test set caratterizzato da un'insorgenza della crisi (*onset*) fissata al campione 800 (cioè minuto 400). Tuttavia, per validare ulteriormente la robustezza del modello nel mondo reale, si è deciso di variare il momento temporale in cui l'evento di stress si manifesta. Infatti, in un contesto clinico reale, le crisi non avvengono secondo uno schema temporale predefinito, anzi sono abbastanza variabili da soggetto in soggetto. Dunque, per scongiurare la probabilità di *overfitting* temporale, sono stati considerati scenari temporali multipli per la validazione del sistema.

Per valutare i due stadi del modello, il codice è stato strutturato per simulare l'anticipo della rampa pre-crisi e del conseguente picco in cinque nuovi scenari temporali: al minuto 150 (campione 300), al minuto 225 (campione 450), al minuto 300 (campione 600), al minuto 350 (campione 700) e al minuto 375 (campione 750).

Inoltre, essendo i dati di validazione generati artificialmente, ogni segnale contiene una forte componente di rumore casuale. Se si fosse testato il modello su una singola traccia per ogni scenario, si sarebbe potuta generare, per pura casualità, un segnale con un rumore molto facile da filtrare, ottenendo prestazioni irrealisticamente alte; oppure, al contrario, una traccia sfortunatamente caotica, che avrebbe fatto fallire il modello. Per scongiurare questo rischio e garantire maggiore solidità al processo di validazione, l'analisi non si è limitata a un solo tentativo ma è stata ripetuta 10 volte per ciascuno scenario temporale. Per implementare ciò è stato usato un ciclo iterativo che ad ogni passaggio modifica il valore di partenza del generatore di numeri pseudocasuali (tramite la funzione `rng`). In sostanza, questa operazione equivale ad aver generato 10 soggetti virtuali diversi per ogni scenario temporale così che tutti avranno l'innesco della crisi esattamente nello stesso minuto, ma ognuno presenterà un rumore di fondo, variazioni cardiache o micromovimenti unici e irripetibili.

Una volta sottoposto l'algoritmo a queste 10 varianti dello stesso evento, sono state estratte un valore medio delle prestazioni e la deviazione standard. Sintomo di elevata

robustezza del modello, a prescindere da come varia il rumore fisiologico di fondo, sono una media alta e una deviazione standard bassa.

8.7.1 Analisi di robustezza temporale del Random Forest

A differenza dei classici approcci basati su soglie statiche, il Random Forest ha dimostrato di non subire alcun degrado prestazionale al variare dell'orario di insorgenza della crisi, restituendo parametri ottimali. Nel dettaglio:

- **Capacità discriminante (AUC):** L'Area Sottesa alla Curva ROC si è mantenuta in un *range* ristretto ed eccellente, compreso tra 0.84 e 0.86, con una deviazione standard estremamente contenuta (oscillante tra ± 0.04 e ± 0.08 sulle 10 ripetizioni). Dal punto di vista statistico, questa costanza certifica che il modello possiede un'elevata capacità di generalizzazione e non è affetto da *overfitting* temporale: l'algoritmo non ha banalmente memorizzato "quando" avviene la crisi nel set di addestramento, ma ha appreso "come" essa si manifesta.
- **Lead Time:** Il tempo di anticipo predittivo ha presentato una media stabile compresa tra i 2.4 e i 2.6 minuti in tutti gli scenari testati. Questo dimostra che il sistema garantisce una finestra d'intervento tempestiva, sia che il sovraccarico fisiologico si presenti all'inizio del monitoraggio, sia che emerga improvvisamente dopo diverse ore di calma.

L'evidenza visiva di questa stabilità è riportata in Figura 8.16. Nel riquadro superiore è possibile osservare come le curve di probabilità del modello si impennino superando la soglia di allarme (25%) in corrispondenza esatta di ogni *onset* simulato. Nei due riquadri inferiori è invece rappresentata graficamente l'invarianza delle metriche principali al variare del minuto della crisi: a sinistra, l'andamento quasi lineare dell'AUC media e, a destra, la costanza dei tempi di anticipo garantiti dall'algoritmo.

Il motivo di questa performance positiva a prescindere dall'arco temporale in cui avviene la crisi, dipende da due aspetti principali: durante la fase di training, si è scelto di rimuovere la variabile temporale assoluta così che l'algoritmo non avesse alcuna informazione su quanto tempo fosse passato dall'inizio della registrazione, ma

è stato forzato a concentrarsi solo su ciò che accade nei segnali istante per istante. Passandogli parametri come la derivata (per il calcolo della pendenza EDA) e la media mobile (per filtrare il rumore), il modello impara a valutare solo le variazioni a breve termine così da evitare il rischio di confondersi con le naturali derive lente del corpo che avvengono nel corso delle ore. L'altro aspetto riguarda la ricerca di condizioni fisiologiche simultanee: infatti, essendo il Random Forest basato su alberi decisionali, lavora incrociando diverse regole e soglie e, non potendo basarsi sul trascorrere dei minuti, ha lanciato l'allarme solo quando si verifica una combinazione specifica di eventi nello stesso istante, come ad esempio un forte aumento dell'EDA accompagnato da un calo improvviso dell'HRV e dell'attività motoria.

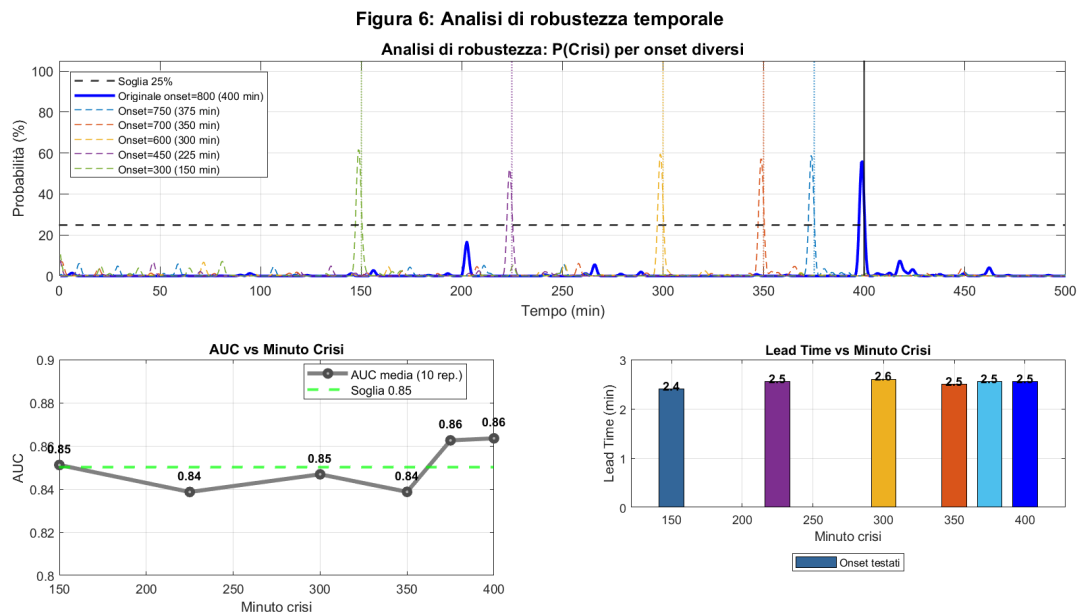


Figura 8.16: Analisi di robustezza temporale del modello Random Forest. In alto: curve di probabilità per i diversi onset testati. In basso: invarianza della capacità discriminante (AUC) e costanza del *lead time* al variare del momento di insorgenza della crisi.

8.7.2 Analisi robustezza temporale della Change Point Analysis

A differenza del Random Forest, l'algoritmo di CPA, si è mostrato abbastanza vulnerabile. Basta osservare il riquadro in basso a destra della Figura 8.17: mentre gli istogrammi blu rappresentanti il tempo di anticipo predittivo garantito dall'RF restano

accettabili e pressoché costanti, la curva rossa, ovvero la percentuale di change point validati, subisce un crollo. Nello scenario originario, con la crisi fissata al minuto 400, la CPA rileva l'evento in tutti i casi, mentre, man mano che la crisi viene anticipata, il tasso di rilevamento precipita fino a toccare lo 0% al minuto 340.

A livello algoritmico, questo è dovuto all'algoritmo di *anomaly detection* utilizzato (*findchangepts*), uno strumento matematico che lavora in modo “globale”, ovvero, considera l'intero tracciato dall'inizio alla fine e cerca il singolo gradino che divide al meglio il segnale in due medie distinte. Infatti, quando la crisi si trova alla fine (ad esempio minuto 400), il segnale risulta essere piatto per ore e presenta un'unica variazione finale e l'algoritmo posiziona il “taglio” esattamente all'inizio della rampa di stress, punto in cui le regole cliniche leggono un'EDA in salita e un HRV in discesa, confermando l'allarme.

Invece, quando la crisi avviene all'inizio (ad esempio minuto 340 o prima), dopo il picco, il corpo comincia a rilassarsi con una fase di discesa molto lenta, protraendosi per oltre cento minuti; l'algoritmo di *anomaly detection*, guardando tutta la registrazione nel suo insieme, a causa di questa enorme discesa, anziché segnalare l'inizio della crisi, posiziona il punto di rottura a metà della fase di rilassamento. A questo punto, il codice passa alle regole cliniche per la validazione e, poiché il soggetto si sta rilassando, la pendenza dell'EDA è negativa; dal momento che il modello è stato addestrato a cercare una pendenza positiva tipica di un attacco imminente, capisce che quel punto è un falso allarme e viene scartato perdendo dunque la vera crisi.

Figura 6: Robustezza temporale del modello ibrido (CPA + RF)

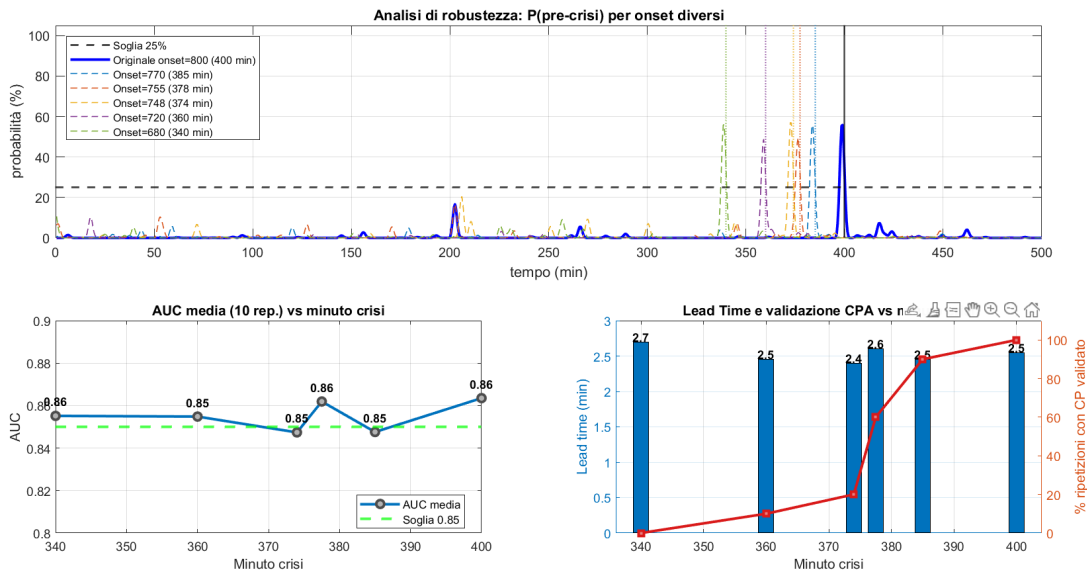


Figura 8.17: Analisi di robustezza temporale del modello ibrido (CPA + RF). Nel riquadro in basso a destra è evidente il crollo della percentuale di *change point* validati (curva rossa) rispetto alla stabilità del *lead time* (istogrammi blu) all'anticipare del minuto di crisi.

8.8 Verso l'applicazione reale: il vantaggio del modello ibrido

Dai risultati emersi dal processo di validazione comprendente l'analisi di robustezza del modello di fronte al rumore (+6 dB) e l'analisi di robustezza alle variazioni temporali della crisi, è stato possibile individuare i punti di forza e i limiti di utilizzo del modello sviluppato.

Sono stati individuati i limiti matematici della CPA di fronte alla variazione temporale dell'inizio della crisi, esaltando però il senso stesso del modello ibrido. Infatti, i due stadi si completano a vicenda colmando le reciproche debolezze:

- La fase di calibrazione:** Il Random Forest non saprebbe gestire i nuovi soggetti se non avesse un riferimento oggettivo della crisi. La CPA, pur essendo vulnerabile se applicata su tracciati di lunga durata, rimane uno strumento efficiente poiché capace di adattarsi in modo automatico e non supervisionato ai dati grezzi di un soggetto ignoto, rilevando il punto di innesco della crisi.

- **La fase predittiva:** Una volta istruito, il RF compensa la rigidità dell’algoritmo matematico. Lavorando esclusivamente sulle variazioni fisiologiche a breve termine, il Random Forest si dimostra immune alle lunghe derive temporali, evita l’*overfitting* e garantisce un *lead time* sicuro e costante in ogni scenario.

Oltre alla sinergia tra i due algoritmi, un altro aspetto rilevante per affacciare il modello al mondo reale è l’applicazione di soglie intra-soggetto. Infatti, tutti i criteri di rilevamento sono calcolati sulla distribuzione statistica individuale del segnale (come l’85° percentile per l’EDA o le mediane per ACC e HRV). In questo modo, l’algoritmo non cerca valori assoluti uguali per tutti (un approccio che fallirebbe data l’enorme eterogeneità fisiologica tipica dello spettro autistico), ma impara a conoscere la “normalità” specifica della persona che indossa il sensore.

Inoltre, il fatto che il modello sia stato addestrato su un soggetto virtuale e testato con successo su un secondo soggetto con parametri fisiologici del tutto diversi (mantenendo intatta la sua finestra predittiva) costituisce un primo riscontro della sua reale capacità di trasferimento tra individui.

Da un punto di vista algoritmico, per una futura implementazione in un dispositivo indossabile operante in *real-time*, sarà sufficiente confinare l’analisi esplorativa della CPA all’interno di finestre temporali mobili (*sliding windows*), affidando la predizione al Random Forest.

In conclusione, l’impiego dei dati artificiali ha rappresentato una base fondamentale per isolare le variabili biologiche e dimostrare la stabilità matematica del modello ibrido in un ambiente controllato. Tuttavia, per quanto simulati con rigore, i dati sintetici non possono replicare interamente l’imprevedibilità del mondo reale. Questo limite fisiologico definisce esattamente le prospettive future e i passaggi necessari per la trasposizione clinica del sistema, che prevedono la validazione dell’architettura su dati reali acquisiti sul campo (ad esempio tramite l’Empatica E4). Poiché i segnali *in vivo* presentano disturbi e artefatti molto più complessi, le soglie cliniche utilizzate finora dovranno essere necessariamente ricalibrate. In termini pratici, pur mantenendo la logica intra-soggetto, si renderà necessario l’uso di tecniche statistiche come la *cross-validazione*: sarà l’algoritmo a fare delle vere e proprie “prove” sui dati reali. Il sistema dividerà i tracciati in vari blocchi e proverà in automatico tante combinazioni

diverse, fino a trovare da solo la configurazione capace di anticipare la crisi con precisione, evitando però di farsi ingannare dalle normali interferenze del sensore o da movimenti casuali del paziente.

Proprio queste prospettive, ovvero le strategie per ricalibrare e generalizzare il modello su pazienti veri, costituiranno il fulcro del prossimo capitolo. L'analisi si sposterà dalla validazione algoritmica in laboratorio alle sfide dell'applicabilità nel mondo reale, culminando con la presentazione di un ulteriore caso di generalizzazione specifico rivolto al monitoraggio dello stress nei piloti, per dimostrare la flessibilità dell'architettura ibrida e la sua capacità di trasformarsi da costruito matematico a strumento preventivo concreto in contesti ad alto carico cognitivo ed emotivo.

9. Generalizzazione del modello ibrido: dai dati sintetici al contesto reale

Sebbene il modello ibrido abbia dimostrato ottime capacità predittive sui dati sintetici generati ad hoc, il passaggio all'utilizzo clinico su pazienti reali tramite dispositivi indossabili commerciali rappresenta una sfida complessa. Si tratta infatti di un vero e proprio "cambio di dominio", poiché si passerebbe da un ambiente controllato ad un contesto reale caotico e imprevedibile.

Come già ampiamente discusso, nei dati artificiali, per quanto disturbati volutamente, i segnali mantengono una certa coerenza di fondo e si comportano secondo regole fisiologiche attese. È doveroso sottolineare che queste regole non sono state ipotizzate a priori, ma sono il frutto diretto della preziosa analisi preliminare condotta sui dati reali: sebbene in quei tracciati mancassero episodi di crisi conclamata, il loro studio, unito alle salde evidenze della letteratura scientifica, è stato fondamentale per comprendere a fondo i *pattern* di base dei soggetti e definire le regole cliniche del modello.

Tuttavia, nel contesto clinico quotidiano, l'imprevedibilità aumenta: i soggetti si muovono bruscamente, hanno variazioni repentine di temperatura corporea, il sensore sfrega sulla pelle o può allentarsi, generando artefatti complessi. Per poter adattare il modello a questo contesto reale, esso manterrà la sua solida architettura ibrida, ma dovrà subire alcune calibrazioni strutturali per gestire le imprevedibilità e le eterogeneità del mondo fisico.

Di seguito verranno presentati i passaggi che si potrebbero implementare per garantire la trasferibilità del modello.

9.1 Pre-elaborazione avanzata: scomposizione dell'EDA e filtraggio artefatti

Come trattato nel Capitolo 8, l'attività elettrodermica è stata elaborata calcolandone semplicemente la media mobile e la derivata per estrarne la pendenza. Questa semplificazione metodologica era matematicamente sufficiente e giustificata dalla natura stessa del dato artificiale: come descritto in precedenza, la componente tonica di base (SCL) era stata simulata tramite una distribuzione di Bates che le garantiva una stabilità controllata nel tempo. Non essendo state simulate le massive e lente derive termiche tipiche della vita reale, la semplice derivata era uno strumento sufficientemente robusto per individuare la rampa pre-crisi senza generare falsi allarmi.

Tuttavia, nel passaggio al monitoraggio reale in vivo, la complessità aumenta drasticamente. Il segnale EDA acquisito sul campo è la somma sovrapposta di due fenomeni distinti: il livello tonico (SCL) e le risposte fasiche (SCR). Nel mondo reale, se un paziente dovesse semplicemente accaldarsi o fare attività fisica, il suo livello tonico di sudorazione salirebbe in modo importante; l'algoritmo predittivo, leggendo questa salita con una semplice derivata, rischierebbe di lanciare un falso allarme, scambiando una normale reazione termica per una rampa di accumulo dello stress.

Per risolvere questo problema è essenziale reintrodurre quel rigore nel *pre-processing* che è stato documentato e testato sui dati reali (si veda il Paragrafo 5.4 del Capitolo 5). L'impiego di algoritmi di ottimizzazione convessa come la *cvxEDA* diventa il passaggio obbligatorio per separare matematicamente la componente legata al sudore termico (SCL) da quella puramente nervosa (SCR).

Solo in questo modo il modello potrà lavorare e calcolare il *change point* esclusivamente sull'attivazione nervosa fasica, garantendo che l'allarme scatti solo per un reale sovraccarico cognitivo ed emotivo. Oltre a questa decomposizione, per i segnali HRV e ACC, sarà fondamentale implementare filtri passa-basso e algoritmi di interpolazione per gestire i vuoti di segnale (NaN) generati dai movimenti del sensore, rendendo il sistema finalmente robusto di fronte alle interferenze.

9.2 Dalle soglie globali alle finestre temporali mobili

Come illustrato nel Capitolo 6 affrontando l'estrazione delle *feature* sui dati reali, è stata già implementata la logica delle finestre temporali mobili (*sliding windows* con *overlap*) per processare il segnale continuo. Tuttavia, nel Capitolo 8, per validare il primo stadio del modello ibrido (ovvero le regole cliniche che confermano il *Change Point*), è stato utilizzato un approccio differente. Si è deciso di impiegare la funzione MATLAB `prctile` per calcolare l'85° o il 75° percentile analizzando l'intera sessione sintetica di 500 minuti in un unico blocco. Questa scelta ha una motivazione ben precisa: prima di affrontare i limiti del tempo reale, era fondamentale assicurarsi che la logica teorica dell'algorithm funzionasse. Se si fossero applicate fin da subito le finestre mobili sui dati sintetici e la predizione avesse fallito, non si sarebbe compreso se l'algorithm sbagliava perché la sua logica matematica era errata, oppure semplicemente perché la finestra temporale era troppo stretta per permettergli di osservare l'arrivo della crisi. Dando all'algorithm una visione totale, è stato possibile eliminare questo dubbio e dimostrare il funzionamento del modello, per poi poterlo adattare alle necessità del mondo reale.

Tuttavia, nella realtà clinica, un dispositivo *wearable* lavora in *real-time* ricevendo i segnali istante per istante e non conosce sin da subito il picco di attivazione dell'intera giornata. A questo limite algoritmico se ne aggiunge uno fisiologico, ovvero, la *baseline* di un essere umano non è piatta, ma fluttua continuamente; se il soggetto sta riposando a letto, avrà parametri vitali molto bassi, mentre, se si alza per fare una passeggiata, il battito e la sudorazione aumenteranno in modo naturale, pur restando in uno stato di totale calma emotiva. Se il sistema applicasse una soglia statica globale calcolata a priori, finirebbe inevitabilmente per scambiare il momento della semplice passeggiata rilassante per un sovraccarico emotivo, generando un falso allarme. Dunque, per far fronte a questa dinamicità fisiologica, l'adeguamento strutturale consisterà nell'estendere la tecnica delle finestre temporali anche nel calcolo delle soglie; in questo modo, anziché usare una percentuale fissa, il codice memorizzerà i dati in un *buffer*

temporaneo per trattenere solo un lasso di tempo recente (ad esempio, gli ultimi 20 minuti di registrazione). Con il passare del tempo, questa finestra avanzerà in modo fluido, scartando i dati vecchi e mantenendo l'osservazione sul presente. Di conseguenza, la logica di validazione cambierà: non si valuterà più se la variabilità dei segnali scende sotto la mediana globale giornaliera, ma ci si chiederà se sta crollando in modo anomalo rispetto a come stava il soggetto negli ultimi 20 minuti. Questo aggiornamento metodologico, che fa uso delle tecniche già esplorate sui dati reali, consentirà al sistema di ricalibrare continuamente il proprio concetto di "calma", senza confondere le normali attività quotidiane con stati di sovraccarico.

9.3 Calibrazione intra-soggetto per l'algoritmo Random Forest

Il modello di RF sviluppato finora è stato istruito per riconoscere il pericolo osservando la salita dei valori fisiologici. Ad esempio, prendendo come riferimento i parametri impostati per la generazione dei dati sintetici nel Capitolo 8, il sistema identifica un rischio elevato quando la sudorazione (EDA) raddoppia rispetto alla sua base, passando ad esempio da 5 a 10 μS . Tuttavia, nella realtà clinica, bisogna tenere conto della fortissima variabilità tra i soggetti, una caratteristica ancora più marcata negli individui con ASD.

Esistono infatti i cosiddetti soggetti "ipo-reattivi", ovvero persone che per natura presentano un'attività elettrodermica cronicamente bassa. Un soggetto di questo tipo potrebbe partire da una base di 1 μS e arrivare al massimo a 4 μS anche durante un momento di stress più acuto. Se si utilizzasse il modello senza correzioni, quel valore di 4 μS verrebbe interpretato come uno stato di assoluta calma, semplicemente perché il numero è basso rispetto alla media della popolazione o ai dati di addestramento. In questo caso, il sistema fallirebbe, non rilevando una crisi reale per quello specifico soggetto.

Dunque, come soluzione si potrebbe procedere con una fase di standardizzazione mediante Z-score; ovvero, prima di dare in input i dati al RF, i numeri grezzi vengono trasformati usando la seguente formula:

$$z = \frac{x - \mu}{\sigma} \quad (9.1)$$

Con questa operazione, al valore x letto dal sensore in quell'istante viene sottratta la media dei valori registrati nell'ultimo periodo (μ) così che il dato viene "pulito" da ciò che per quel soggetto è la normalità. Il risultato viene poi diviso per la deviazione standard (σ) che rappresenta quanto il segnale di quella specifica persona tende a oscillare normalmente. Grazie a questo processo, l'algoritmo smette di guardare i numeri "fissi" e inizia a ragionare in termini di intensità relativa, osservando di quanto è aumentato il segnale rispetto alla media di questa persona. In questo modo si può comprendere se si tratta di un parametro preoccupante rispetto alla storia recente del soggetto. Questo permette allo stesso identico modello di funzionare perfettamente su chiunque: sia su chi ha reazioni fisiologiche molto forti, sia su chi manifesta lo stress in modo molto più silenzioso e lieve.

9.4 Monitoraggio *real-time* e elaborazione *streaming* dei dati

Per avere un sistema di elaborazione in *streaming*, l'architettura ideata deve subire alcuni cambiamenti:

- **Architettura ad eventi:** a differenza dello *script* utilizzato nei test, che leggeva i dati in modo sequenziale da un file già salvato, il sistema reale deve operare in modo reattivo. Per ottimizzare il consumo energetico dello *smartwatch*, il *software* non resterà costantemente in esecuzione, ma passerà la maggior parte del tempo in uno stato di sospensione. Il codice si attiverà solo nel momento esatto in cui il modulo *Bluetooth* del braccialetto invierà un nuovo pacchetto di dati. Questa gestione permette al sistema di elaborare le informazioni in tempo reale solo quando necessario, garantendo una lunga durata della batteria del dispositivo senza perdere la tempestività dell'analisi.
- **Gestione della memoria:** uno *smartwatch* non ha la capacità di memoria di un computer; quindi, non può tenere in linea ore e ore di dati. Dunque, il si-

stema dovrebbe usare una memoria a breve termine chiamata "buffer circolare". Ogni volta che arriva un nuovo dato dal braccialetto, questo prende posto sul *buffer* e, contemporaneamente, il dato più vecchio va fuori per lasciargli spazio. Dal punto di vista informatico, questa struttura si implementa allocando un *array* di dimensione fissa (N) e utilizzando l'operatore matematico modulo per aggiornare l'indice di scrittura: $i = (i \bmod N) + 1$. Questo stratagemma permette all'indice di tornare automaticamente a zero una volta raggiunta la fine dell'*array*, sovrascrivendo fisicamente il dato più vecchio senza la necessità di traslare gli altri valori in memoria. In questo modo, il dispositivo non si ingolfa mai con troppe informazioni e ha sempre a disposizione solo i dati più nuovi e aggiornati. Per garantire la reattività del sistema in *streaming*, l'elaborazione del *buffer* avverrebbe con una frequenza di aggiornamento coerente con quella del sensore, assicurando che ogni nuovo campione venga processato senza ritardi percettibili.

- **Esecuzione mirata della CPA:** Applicando l'algoritmo `findchangepts` esclusivamente a questo piccolo *buffer*, il costo di calcolo si riduce drasticamente. Il sistema deve solo verificare se, in quei pochi minuti di memoria recente, l'EDA ha cambiato inclinazione.
- **La validazione:** Se la CPA rileva un'anomalia nel *buffer*, innesca la validazione clinica. Solo se si verifica la contemporaneità tra sbalzo dell'EDA, crollo motorio e ritiro vagale, i dati vengono passati al *Random Forest* per l'emissione dell'allarme.

Bisogna tenere conto anche dei primissimi minuti in cui il soggetto indossa il braccialetto, in cui il *buffer* risulta vuoto o con troppi pochi dati per effettuare calcoli affidabili. Inoltre, vi è un altro limite legato però all'hardware, già documentato nel Capitolo 5: quando si allaccia il dispositivo, i sensori necessitano di un certo tempo di assestamento affinché si crei il contatto ideale tra elettrodi e pelle. Se il sistema cominciasse a valutare le anomalie sin dal primo secondo, interpreterebbe questa instabilità iniziale come una crisi imminente, generando falsi allarmi. Per risolvere queste problematiche, il *software* potrebbe prevedere un periodo di tolleranza iniziale, ad esempio di

15 minuti, fase in cui il dispositivo si limiterà a immagazzinare i dati senza attivare allarmi; solo quando i sensori hanno raggiunto stabilità e vi saranno dati attendibili a sufficienza, il sistema potrà emettere valutazioni di rischio.

Questo processo di transizione permetterebbe dunque di avere un *software* leggero, reattivo e capace di garantire al *caregiver* un tempo di preavviso necessario per intervenire tempestivamente.

10. Validazione cross-domain: generalizzazione del modello predittivo sui piloti aerei

L'obiettivo finale di un modello di *Machine Learning* applicato alla bioingegneria non è solo quello di funzionare in un ambiente controllato, ma di dimostrare che le logiche matematiche su cui si fonda possono essere generalizzate e applicate in contesti universali. Dunque, per dimostrare la robustezza e la trasferibilità del modello ibrido sviluppato, si è deciso di sottoporlo ad un test estremo di generalizzazione, ovvero, trasferire il modello dalla predizione del sovraccarico emotivo negli ASD alla predizione dello stress in piloti d'aereo. Naturalmente, variando il dominio e il contesto applicativo, è stata necessaria una profonda ricalibrazione fisiologica e algoritmica, guidata da fondamenta teoriche e implementative descritte di seguito.

10.1 Basi fisiologiche e revisione della letteratura

Il primo problema da cui si è dovuti partire è che la risposta fisiologica allo stress in un adulto sano è strutturalmente e temporalmente diversa dalla crisi comportamentale di un soggetto con ASD. Analizzando la letteratura scientifica recente, sono emersi tre aspetti fondamentali per la ricalibrazione del modello:

- **Dinamica temporale:** Spesso lo stress viene trattato come una entità statica, ma come ampiamente dimostrato da Rosenblum et al.[43], lo stress cognitivo non si manifesta come un picco esplosivo, ma come un accumulo progressivo, che nel pilota si traduce in un *plateau* sostenuto di 10–15 minuti [44]. Di conseguenza, il classificatore non deve più cercare un incremento repentino, ma una rampa lenta e prolungata.
- **Integrazione multi-sensore e attivazione simpatica:** I piloti possiedono un'eccellente flessibilità autonoma, con una *baseline* dell'attività elettrodermica molto contenuta. Tuttavia, Coelli et al. [41] hanno dimostrato che l'ampiezza

fasica dell'EDA subisce un aumento netto sotto stress mentale, mentre la variabilità cardiaca (HRV RMSSD) subisce un calo graduale che risulta però ritardato rispetto all'attivazione simpatica [42]. Per catturare istantaneamente l'innesco dello stress, come suggerito da Wang et al. [44], è stato fondamentale introdurre una quinta *feature* nel sistema: la frequenza cardiaca (HR), che rappresenta il parametro di attivazione più rapido, anticipando il crollo dell'HRV.

- **Il ruolo dell'accelerometro:** Nel contesto dell'autismo, l'accelerometro mappa i movimenti di *stimming* autoregolatori. Ma, all'interno del *cockpit*, il corpo del pilota è ancorato al sedile e fisicamente inattivo. Ciononostante, Yu et al. [45] hanno evidenziato che le accelerazioni traslazionali passive (tipiche del volo) agiscono come *stressor* fisici diretti in grado di innescare risposte dell'EDA. Questo ha confermato l'importanza di mantenere l'accelerometro attivo nel processo di monitoraggio multimodale.

Tale ricalibrazione non è stata però solo qualitativa, ma ha richiesto una precisa quantificazione dei range fisiologici attesi, al fine di modellizzare correttamente il passaggio dalla disregolazione emotiva acuta dell'ASD al carico cognitivo operativo del pilota.

10.2 Generazione del dataset fisiologico

In assenza di un *dataset* clinico accessibile contenente registrazioni di volo complete, si è proceduto alla generazione di una nuova matrice di dati sintetici. Il *dataset* simula il monitoraggio di 10 piloti virtuali ($N = 10$), ciascuno osservato per 500 minuti. Per simulare i vincoli di memoria e trasmissione di un reale dispositivo indossabile, la frequenza di campionamento è stata impostata a un dato ogni 30 secondi ($f_s = 0.5$ min).

Le variabili predittive sono state aumentate a cinque: EDA tonica/fasica, *Slope* dell'EDA, HRV, accelerometria (ACC) e frequenza cardiaca (HR). Per garantire il massimo realismo, a ciascun segnale è stato applicato un rumore stocastico per simulare le naturali fluttuazioni fisiologiche di base (ad esempio, un HR a riposo di circa 75 bpm e un'accelerazione di gravità stabile a 1.0 g).

La scelta dei parametri per la generazione di tali dati riflette la profonda differenza fisiologica tra la disregolazione emotiva acuta in soggetti con ASD e il sovraccarico cognitivo o cinemático di un pilota in assetto operativo. Analizzando i dati estratti dalla letteratura, i valori di riferimento sono stati ricalibrati come riassunto nella Tabella 10.1.

Tabella 10.1: Calibrazione dei parametri fisiologici (Dominio ASD vs. Dominio Piloti)

Parametro	Crisi ASD (Cap. 8)	Stress Piloti (Cap. 10)
EDA baseline	1–4 μS	0.5–2 μS (più bassa)
EDA picco evento	+8–10 μS	+2–4 μS (più contenuta)
HRV RMSSD baseline	~50 ms	55–70 ms (piloti allenati)
HRV calo stress	–50% (crollo)	–20–30% (graduale)
ACC durante evento	Crollo motorio (0.4x)	Minima variazione (seduti)
Durata fase pre-evento	3 min (rampa esplosiva)	10–15 min (accumulo lento)
Forma del segnale	Picco acuto	<i>Plateau</i> sostenuto

I valori presenti in tabella possono essere giustificati e modellizzati per l’addestramento dell’algoritmo nel modo seguente:

- **EDA baseline (0.5–2 μS) e picco evento (+2–4 μS):** Nel contesto del volo, l’attività fisica è confinata a micromovimenti in un ambiente molto controllato. Di conseguenza, il livello tonico di base (SCL) parte da valori più bassi. Quando si verifica un evento di stress operativo, la risposta simpatica fasica si attiva rapidamente [45]. Tuttavia, il pilota è addestrato a reprimere la reazione emotiva. Questo si traduce in un segnale EDA che mostra attivazioni frequenti ma di ampiezza moderata, lontane dai valori elevati di +8–10 μS tipici di un sovraccarico autistico.
- **HRV RMSSD Baseline (~55–70 ms) e calo durante lo stress (–20–30 %):** I piloti sono adulti sani con un’eccellente flessibilità autonoma. Sotto forte stress, il sistema parasimpatico si ritira, ma nei piloti questo ritiro non assume i contorni di un crollo vertiginoso, bensì di un adattamento progressivo. Come suggerito

da Rosenblum et al.[43], l'attivazione simpatica (aumento HR) è istantanea, ma la riduzione dell'HRV è fisiologicamente ritardata, diventando significativa solo nella fase centrale dell'esposizione allo stress.

- **ACC durante l'evento (minima variazione):** In una crisi ASD si registra frequentemente un'esplosione di agitazione motoria. Nel *cockpit*, pur subendo intensi carichi mentali e forze traslazionali, il corpo del pilota è saldamente ancorato al sedile. L'accelerometro misurerà quindi solo le vibrazioni del velivolo e i micromovimenti finalistici legati al pilotaggio.

L'inclusione di queste caratteristiche ha permesso di generare i dati seguendo due fasi distinte:

- **Fase di Pre-Stress:** Lo stress del pilota cresce parallelamente all'evoluzione della fase critica del volo (es. avvicinamento e atterraggio). L'accumulo del carico cognitivo è stato simulato con un incremento lineare dell'EDA della durata di 10–15 minuti, accompagnato da una salita rapida dell'HR e da un calo progressivo dell'HRV. Questa dinamica produce un *plateau* di allerta prolungata.
- **Fase di Recovery:** Per imitare il ritorno all'omeostasi [43], il decadimento dei parametri nella fase post-crisi ha seguito una curva esponenziale, rispecchiando la lentezza fisiologica del sistema nervoso nel riassorbire i marcatori dello stress.

10.3 Evoluzione dell'architettura ibrida

Per gestire questo nuovo dominio, l'algoritmo a due stadi è stato riadattato con il duplice obiettivo di garantire la massima efficacia in condizioni di forte sbilanciamento dei dati e operare in tempo reale, senza poter osservare il futuro.

10.3.1 Stadio 1: CPA e baseline personalizzate

Il primo stadio mantiene l'utilizzo della *Change Point Analysis* (CPA) per individuare le rotture strutturali nel segnale. Data l'altissima variabilità interindividuale nella gestione dello stress aeronautico [44], le soglie sono state rese strettamente individuali.

In termini pratici, in un contesto reale, il sistema sfrutta una finestra di calibrazione pre-volo (ad esempio i primi 15–30 minuti durante i controlli nel *cockpit*) per stabilire i percentili base di quello specifico pilota. Successivamente, l'evento viene validato clinicamente solo se l'EDA supera l'85° percentile della propria *baseline* e l'HRV scende sotto la propria mediana. In questo modo, la CPA ha ottenuto una validazione positiva su tutti i soggetti simulati.

10.3.2 Stadio 2: Dal Random Forest al RUSBoost

Nel modello iniziale applicato all'ASD, il secondo stadio, responsabile della predizione, era costituito da un *Random Forest* basato su logica *bagging*; nel dominio ASD, per via della presenza di fluttuazioni emotive continue e imprevedibili, il RF si è mostrato robusto; tuttavia, nonostante ciò, esso presenta un limite di fronte a *dataset* molto sbilanciati, ovvero tende ad assecondare la classe maggioritaria ignorando quella minoritaria. Nel dominio aeronautico, la finestra di pre-stress rappresenta una frazione molto piccola del volo quindi, l'utilizzo di un RF avrebbe causato un crollo della sensibilità.

Per questo motivo, l'architettura è transitata verso un modello RUSBoost. La differenza fondamentale tra i due algoritmi risiede sia nel modo in cui gestiscono i dati, sia nella loro architettura di apprendimento: mentre il RF è un modello "parallelo" (logica *bagging*), il RUSBoost è un modello "sequenziale" che unisce due tecniche specifiche per i dati sbilanciati:

1. **Random Under-Sampling:** Il RUSBoost, prima di iniziare a studiare, esegue un "sottocampionamento stocastico", ovvero scarta in modo casuale una grandissima quantità di dati appartenenti alla classe di maggioranza (la calma), riducendone il volume fino a raggiungere un rapporto matematico imposto di 3 a 1 rispetto allo stress. In questo modo, quel piccolo momento di stress diventa visibile e rilevante.
2. **Boosting Sequenziale:** Nel *Random Forest*, vengono creati (ad esempio) 100 alberi decisionali che lavorano in parallelo, in modo del tutto indipendente l'uno dall'altro; alla fine, si limitano a fare una votazione di maggioranza. Il RUSBoost invece addestra gli alberi in modo sequenziale: il primo albero analizza i dati e

inevitabilmente commette degli errori di previsione. A questo punto, l'algoritmo individua i campioni classificati male, assegna loro un peso matematico maggiore e li passa al secondo albero. L'albero successivo è così costretto a concentrarsi proprio sugli errori del suo predecessore per correggerli.

Questi due processi combinati si rivelano fondamentali: il bilanciamento impedisce allo stress di essere celato dal rumore di fondo, mentre il *boosting* costringe il modello a specializzarsi nell'identificazione dei precursori fisiologici dello stress, aumentando la sensibilità e contrastando il rischio di mancati allarmi.

10.4 L'elaborazione in tempo reale: masking e filtro causale

L'implementazione dello stadio 2 ha richiesto due complessi interventi strutturali sui dati, finalizzati a rendere il modello applicabile al contesto aereo:

1. **Prevenzione del *data leakage*:** Da un punto di vista fisiologico, con la fine logica di un'emergenza non vuol dire che si ha un ritorno immediato dei parametri alla normalità. Se in fase di addestramento l'algoritmo avesse osservato la fase di lento decadimento post-crisi (etichettata come calma), avrebbe subito un grave inquinamento dei dati (*data leakage*). Il modello avrebbe infatti imparato ad associare valori elevatissimi di sudorazione e un'alta frequenza cardiaca a uno stato di tranquillità, falsando la propria capacità predittiva. Per impedire questo errore metodologico, è stata implementata una maschera logica (*masking*). Questa è stata realizzata tramite indicizzazione logica: il codice genera un vettore che assegna il valore "falso" a tutti i campioni temporali successivi all'inizio della crisi, escludendoli fisicamente dal *set* di addestramento. In questo modo il RUSBoost studia esclusivamente la rampa fisiologica che porta dalla calma totale al pre-stress.
2. **Filtro causale e soglia decisionale:** Gran parte dei modelli predittivi in letteratura ottiene precisioni elevatissime sfruttando filtri di *smoothing* simmetrici

(ovvero algoritmi che smussano la curva guardando contemporaneamente sia al passato che al futuro del segnale). Tuttavia, in un *cockpit*, il computer di bordo non possiede i dati del futuro. Dunque, il risultato in uscita dal RUSBoost è stato livellato utilizzando un filtro asimmetrico: il sistema calcola la media mobile esclusivamente sugli ultimi 14 campioni acquisiti (i 7 minuti di storia passata), guardando a zero campioni nel futuro, garantendo così un funzionamento *Real-Time* matematicamente realistico.

Inoltre, gran parte dei classificatori restituisce un risultato sotto forma di probabilità percentuale (da 0 a 100%). In queste architetture, chi progetta è costretto a scegliere una soglia di allarme in modo arbitrario, introducendo un fattore di soggettività umana che può compromettere l'affidabilità del sistema. Invece, il RUSBoost non ragiona in percentuali ma sostituisce un *classification score* continuo, cioè un punteggio numerico che può essere sia negativo che positivo. I valori negativi indicano che prevalgono le evidenze fisiologiche a favore dello stato di calma, i valori positivi indicano invece che i segnali legati allo stress hanno preso il sopravvento. Di conseguenza, la soglia di innesco dell'allarme non deve essere stabilita per tentativi, ma è ancorata in modo naturale e rigoroso allo zero. L'esatto istante in cui la curva del punteggio passa da negativa a positiva certifica oggettivamente e matematicamente il passaggio allo stato di pericolo.

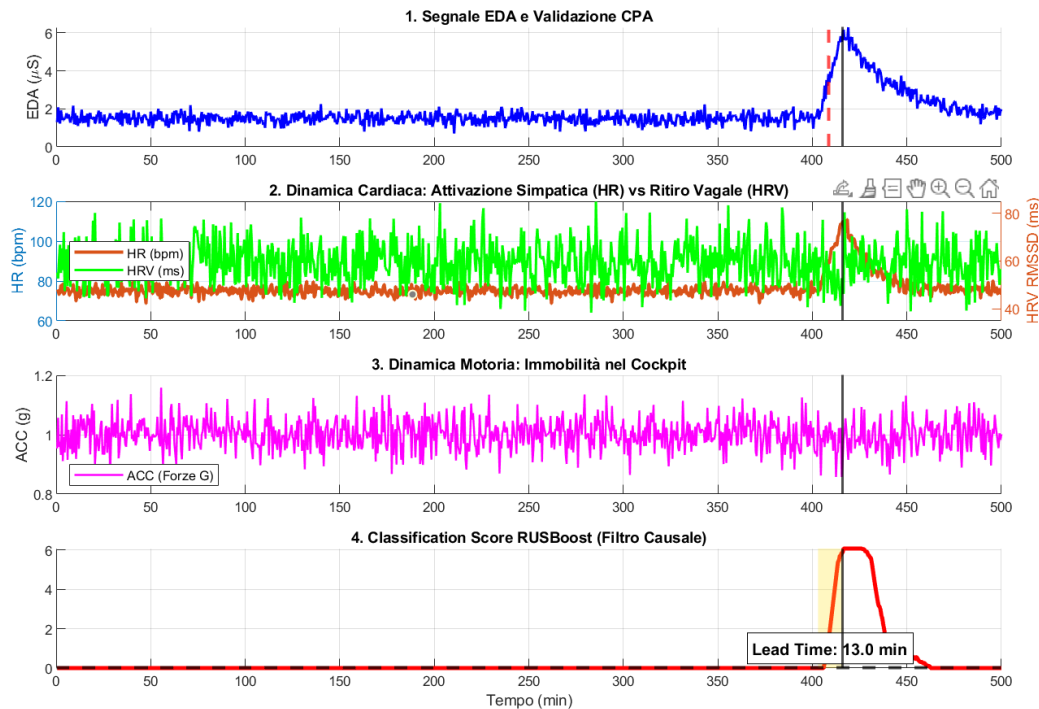


Figura 10.1: Sintesi vista dell'architettura ibrida applicata a un soggetto di test estratto dalla LOO-CV. I quattro pannelli illustrano l'evoluzione temporale dell'evento di stress cognitivo: (1) **Segnale EDA**: La linea rossa tratteggiata indica l'innesco dell'*allostatic load* individuato autonomamente dalla *Change Point Analysis* (Stadio 1) e validato clinicamente dalla *baseline*. (2) **Dinamica cardiaca**: Evidenzia l'attivazione simpatica istantanea (salita dell'HR, in marrone) contrapposta al lento ritiro vagale (calo dell'HRV, in verde). (3) **Dinamica motoria**: I valori dell'accelerometro certificano l'assenza di artefatti da movimento, tipica dell'ancoraggio del pilota al sedile. (4) **Classification Score (Stadio 2)**: La curva di *output* del modello RUSBoost, livellata dal filtro causale [14 0], attraversa la soglia dello zero matematico. L'area gialla evidenzia l'*Early Warning* generato dal sistema, garantendo in questo specifico volo un *lead time* operativo di 13.0 minuti prima dell'apice della crisi (linea nera continua).

10.5 Validazione e risultati

Per valutare la robustezza inter-soggetto, è stata implementata una validazione incrociata di tipo *leave-one-out* (LOO-CV) a 10 fold; cioè l'algoritmo viene addestrato iterativamente sui dati di 9 piloti e testato sul decimo, garantendo che il modello venga valutato esclusivamente su tracciati fisiologici che non ha mai visto in fase di *training*.

I risultati aggregati (riassunti nella Curva ROC e nelle Matrici di Confusione in Figura 10.2 e 10.3) riassumono un valore di alto livello:

- **AUC (Area Under Curve)**: 0.89 ± 0.04
- **Sensibilità (Recall)**: 79.0%

- **Specificità:** 99.1%
- **Lead Time Medio:** 10.1 minuti

Il conseguimento di una specificità quasi perfetta (99.1%) certifica l'efficacia combinata del *masking* e del filtro causale nell'abbattere i falsi positivi durante le lunghe ore di volo tranquillo. Parallelamente, un'AUC di 0.89 e una Sensibilità del 79.0% rappresentano valori di eccellenza, posizionandosi nei *range* migliori presenti in letteratura per il rilevamento dello stress tramite dispositivi indossabili [42]. Anche un tempo di preavviso di oltre 10 minuti fornisce un margine temporale sufficiente per attuare in sicurezza le *Standard Operating Procedures* (SOP) di mitigazione del rischio (ad esempio, l'inserimento preventivo del pilota automatico, il passaggio dei comandi al co-pilota o l'esecuzione di tecniche di respirazione prima di manovre ad alto carico).

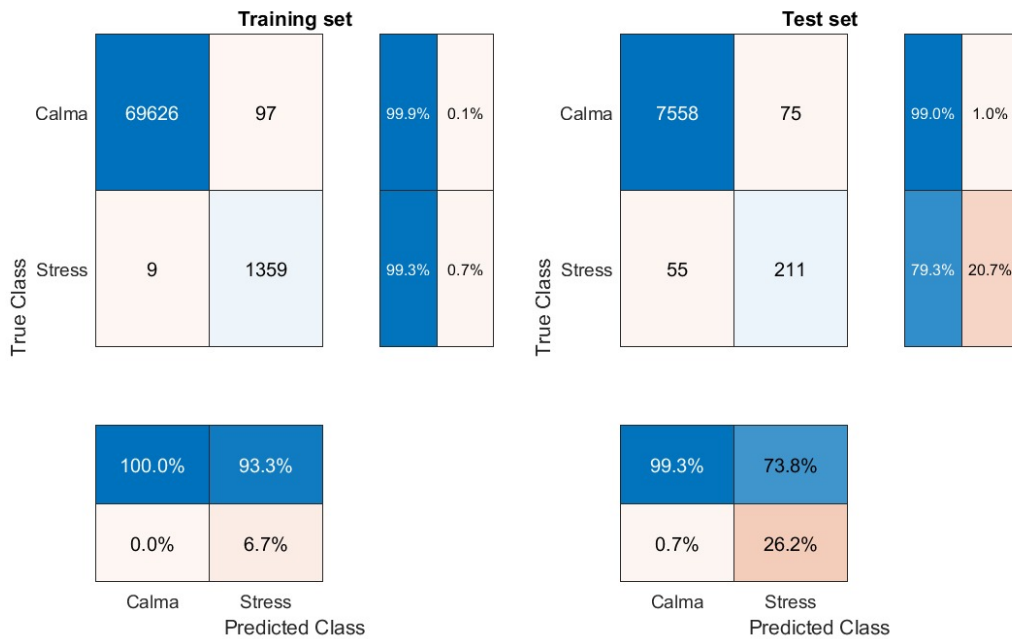


Figura 10.2: Matrici di Confusione generate dall'algorithmo RUSBoost. Da sinistra verso destra: prestazioni sul set di training, prestazioni sul set di test (LOO-CV)

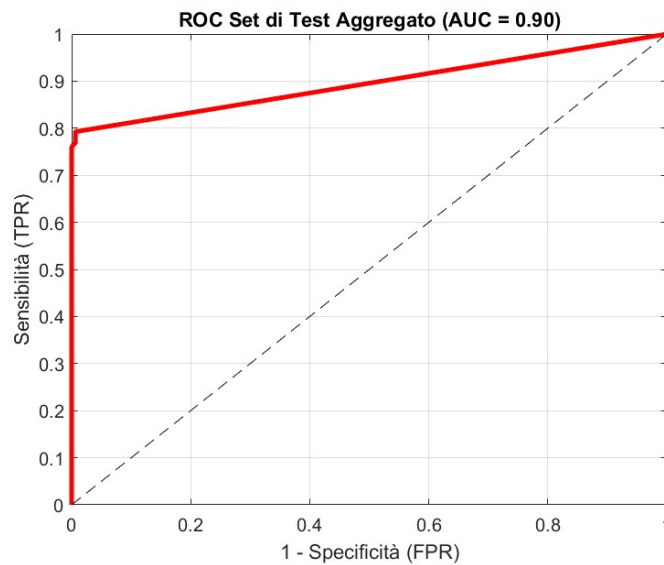


Figura 10.3: Curva ROC (calcolata sul set di test aggregato).

10.5.1 Analisi dell'importanza predittiva e il ruolo dell'integrazione multimodale

L'analisi dell'importanza predittiva delle *feature* estratta dall'*Ensemble* RUSBoost ha rivelato dinamiche decisionali molto nette. Come si evince dalla Figura 10.4, la frequenza cardiaca (HR) domina la classificazione. Immediatamente a ridosso dell'HR, il modello ha assegnato un peso predittivo fondamentale ai valori assoluti dell'attività elettrodermica (EDA), confermandola come indicatore strutturale dello stress.

Tuttavia, dal grafico si osserva anche il totale azzeramento del peso relativo all'HRV, allo *Slope* dell'EDA e all'Accelerometro (ACC). Questo annullamento non deriva da un'inutilità intrinseca di queste variabili, ma rappresenta un artefatto matematico indotto dalla natura sintetica dei dati combinata con l'efficienza estrema degli alberi decisionali. Il RUSBoost, essendo un algoritmo di ottimizzazione orientato alla massima separazione delle classi, tende a sfruttare le variabili più "pulite" e dirette. In presenza di segnali simulati privi di gravi interferenze meccaniche, il modello ha ritenuto sufficiente la diade HR-EDA per massimizzare le prestazioni, ignorando i canali più complessi o lenti.

Naturalmente, nel contesto di un volo reale, questo scenario ideale non esisterebbe. Solo l'HR e l'EDA non sarebbero sufficienti, poiché fattori di confondimento

ambientali (come variazioni microclimatiche) o sforzi fisici imprevisti ne abbasserebbero l'affidabilità esclusiva, esponendo il sistema al rischio di continui falsi positivi. Nel mondo reale, l'inevitabile "rumore" fisiologico che corrompe i sensori principali costringerebbe l'algoritmo a ripescare attivamente il valore latente dell'HRV e le componenti derivative come lo *Slope* per confermare il sovraccarico cognitivo profondo e scartare gli artefatti. Questo dimostra che l'architettura a 5 *feature* non rappresenta solo un'opzione preferibile, ma costituisce un requisito bioingegneristico essenziale per garantire l'affidabilità clinica del modello finale.

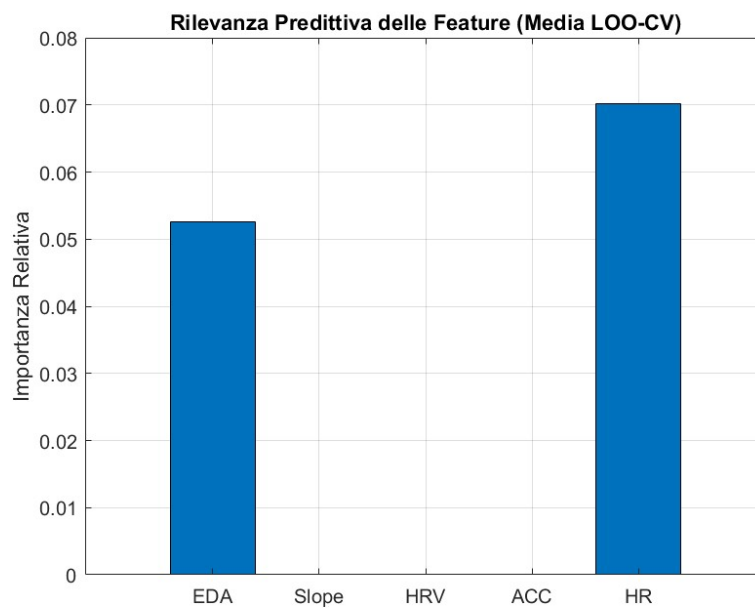


Figura 10.4: Rilevanza predittiva delle *feature* calcolata come media sulla *leave-one-out cross-validation* (LOO-CV). L'istogramma evidenzia come l'algoritmo RUSBoost, operando su dati sintetici ideali, attribuisca il massimo potere discriminante alla frequenza cardiaca (HR) e all'attività elettrodermica (EDA).

10.5.2 Prestazioni e robustezza inter-soggetto

Per verificare l'assenza di *overfitting*, è stato eseguito un confronto diretto tra l'evoluzione dello *score* di classificazione calcolato su un soggetto appartenente al *set* di addestramento (Dati Noti) e quello calcolato su un soggetto isolato nel *set* di validazione (Dati Ignoti). Come illustrato nella Figura 10.5, le due curve mostrano un andamento quasi uguale: in entrambi i casi, l'algoritmo mantiene uno *score* sullo zero durante le ore di volo tranquillo, per poi innalzarsi bruscamente generando l'allarme solo in con-

comitanza dell'accumulo del carico cognitivo. Il fatto che la curva rossa (Dati Ignoti) replichi la precisione e la pulizia della curva blu (Dati Noti) è la dimostrazione matematica che l'algoritmo RUSBoost ha appreso la regola fisiologica universale dello stress, non i semplici *pattern* del *set* di addestramento.

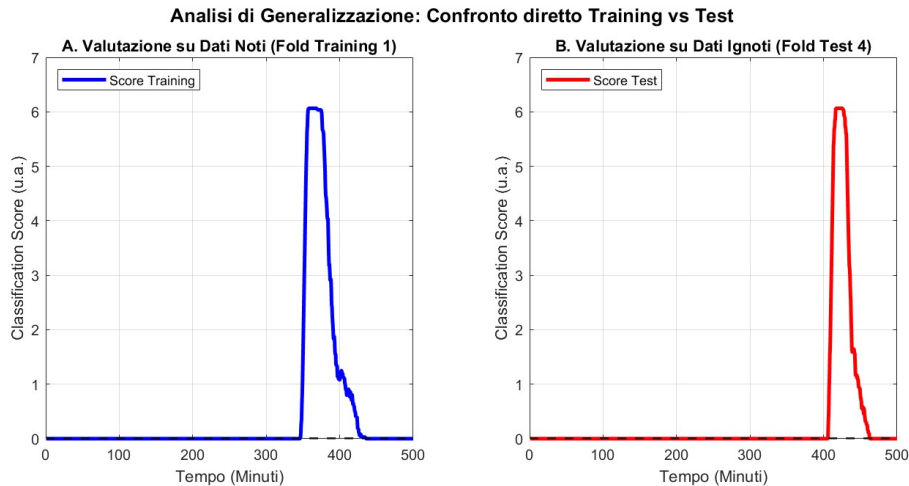


Figura 10.5: Analisi di generalizzazione. Il confronto diretto tra l'andamento del *Classification Score* su un soggetto del *set* di addestramento (A) e su un soggetto ignoto del *set* di test (B) evidenzia una morfologia identica. Il mantenimento della stabilità in fase di calma e la reattività all'innesco dello stress sui dati ignoti confermano l'assenza di *overfitting*.

L'analisi è stata poi estesa all'intera popolazione per valutare la robustezza inter-soggetto. Come mostrato nel pannello superiore della Figura 10.6 (Dinamica dello *Score* di Classificazione per l'intera popolazione), il modello si comporta in modo estremamente coerente su tutti e 10 i piloti simulati: per nessuno di essi si registrano falsi positivi (picchi anomali) durante la lunga fase di volo stabile iniziale. Ogni singola curva si innesca esclusivamente nella finestra temporale in cui è stato generato l'evento critico (tra il minuto 350 e 450).

Questa costanza è ulteriormente confermata dai seguenti valori individuali:

- **AUC per soggetto:** La capacità di discriminazione rimane elevata, oscillando costantemente intorno alla media di 0.89 (linea tratteggiata rossa) senza alcun crollo prestazionale critico al di sotto di 0.85.
- **Lead time per soggetto:** Il tempo di preavviso si distribuisce in modo omogeneo intorno alla media di 10.1 minuti. Pur in presenza della fisiologica variabilità interindividuale (con tempi di reazione che spaziano da circa 8 a 13 minuti),

il sistema garantisce per ogni singolo volo un margine temporale ampiamente sufficiente per attivare le manovre di sicurezza in cabina.

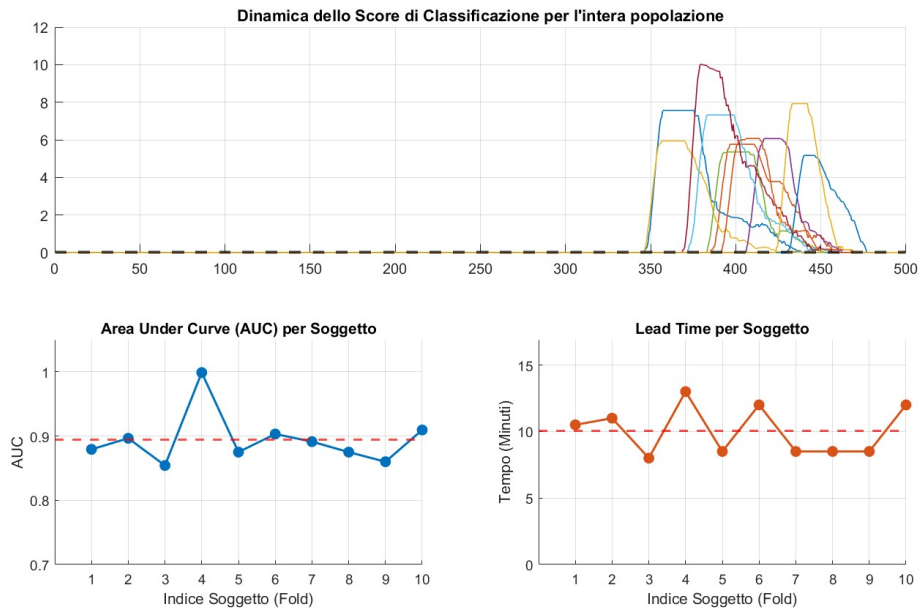


Figura 10.6: Analisi della robustezza inter-soggetto sull'intera popolazione ($N = 10$). Il pannello superiore mostra la coerenza temporale dell'innescio degli allarmi (assenza totale di falsi positivi pre-crisi). I grafici inferiori dimostrano la stabilità delle prestazioni individuali, con valori di AUC e *Lead Time* (minuti) costantemente allineati o superiori ai requisiti operativi medi (linee tratteggiate rosse).

10.6 Confronto tra il modello clinico e il modello aeronautico

Il passaggio dal dominio clinico (soggetti con ASD) al dominio aeronautico non si è limitato a una semplice ricalibrazione delle soglie fisiologiche, ma ha comportato veri e propri cambiamenti dell'architettura per adattarla al nuovo contesto e alle nuove necessità per un corretto monitoraggio.

Di seguito vengono elencati i principali cambiamenti rispetto al modello originario del presente lavoro:

1. **Integrazione multimodale:** Il modello ASD utilizzava 4 *feature* per la predizione, escludendo la frequenza cardiaca (HR) perché ritenuta meno informativa

anche a causa degli inquinamenti dovuti alla possibile e intensa attività motoria dei soggetti. Nel modello aeronautico, l'immobilità del pilota ha permesso di innalzare le *feature* a 5, reintegrando l'HR. Questa scelta si è rivelata decisiva per mappare l'attivazione simpatica istantanea, che agisce in modo complementare e anticipato rispetto al più lento ritiro vagale [44].

2. **Validazione:** Il modello originale era stato testato su un *set* statico e numericamente limitato ($N = 2$) a causa della reale difficoltà di acquisire dati in quantità sufficiente e “puliti” per questi soggetti. Per il dominio aereo, la validazione è transitata verso una *Leave-One-Out Cross-Validation* (LOO-CV) su 10 soggetti; un approccio che permette di validare la capacità dell'algoritmo di generalizzare lo stress su individui che non ha mai “visto” in fase di addestramento.
3. **Prevenzione del data leakage:** Nel primo modello, l'algoritmo osservava l'intero vettore temporale. Nel modello per piloti, è stato introdotto il *Masking* logico per “nascondere” all'algoritmo la fase di recupero (*recovery*). Questo accorgimento impedisce al sistema di fare confusione su stati fisiologici ambigui (ad esempio, quando il pilota è ormai psicologicamente rilassato ma la sua pelle continua a presentare un'elevata sudorazione residua).
4. **Causalità Temporale:** Gran parte delle validazioni cliniche “da laboratorio” sfrutta filtri simmetrici, che guardano simultaneamente sia ai dati passati che a quelli futuri per smussare le curve. Per garantire l'applicabilità in volo, il nuovo modello utilizza esclusivamente una media mobile causale asimmetrica (14 campioni passati, 0 futuri), simulando le reali limitazioni di un computer di bordo.
5. **Criterio decisionale:** Il classificatore precedente si basava su una soglia di probabilità empirica (scelta arbitrariamente intorno al 25%). Il passaggio all'algoritmo RUSBoost ha permesso di utilizzare un *Classification Score* continuo, che ancora la decisione di innescare l'allarme in modo rigoroso e naturale allo zero matematico, eliminando ogni soggettività umana.

11. Conclusioni e prospettive future

Il punto di partenza di questo lavoro è stata una problematica clinica importante: dare voce a chi non può verbalizzare il proprio disagio. In soggetti con autismo a basso funzionamento (MV-ASD), la crisi comportamentale non è quasi mai un evento improvviso e immotivato, ma l'esplosione finale di un sovraccarico emotivo rimasto a lungo invisibile agli occhi di chi osserva. Attraverso l'uso dei sensori *wearable* e lo sviluppo di un modello ibrido di *Machine Learning*, si è provato a gestire questa silente fisiologia, trasformando i segnali biologici del corpo in uno strumento preventivo di allarme.

11.1 Il significato clinico dei risultati ottenuti

Il lavoro è partito dall'analisi di dati reali raccolti in un ambiente multisensoriale, dove si è appreso che la qualità del dato dipende strettamente dal contesto. L'attività elettrodermica (EDA) e la variabilità cardiaca (HRV) si sono rivelate ben più che semplici parametri numerici: esse tracciano in tempo reale il delicato equilibrio tra l'eccitazione del sistema simpatico, che agisce da acceleratore, e l'azione frenante del nervo vago.

Il risultato più significativo dell'architettura ibrida (CPA + Random Forest) è stata la dimostrazione che è possibile ottenere un tempo di anticipo compreso tra 1.5 e 3 minuti. In un contesto clinico o domestico, questi minuti rappresentano il tempo vitale necessario a un educatore o a un genitore per intervenire, cambiare stimolo o attuare precise manovre calmanti, evitando che il sovraccarico degeneri in una crisi aggressiva.

La validazione *cross-domain* effettuata sui piloti ha inoltre confermato la robustezza del modello e la sua capacità di generalizzazione: sebbene i parametri siano stati ricalibrati per un diverso livello di stress, la logica matematica di base ha dimostrato di saper riconoscere l'accumulo di tensione a prescindere dal soggetto, confermando la flessibilità del sistema.

11.2 I limiti del lavoro

Nonostante l'efficacia del modello, è fondamentale riconoscere i limiti che hanno caratterizzato questa ricerca, necessari per inquadrare correttamente i risultati:

- **Assenza di un dataset di validazione clinica reale:** L'impossibilità di disporre di un dataset clinico reale sufficientemente ampio e annotato con crisi esplosive ha reso obbligatorio il ricorso ai dati sintetici. Sebbene i parametri siano stati calibrati con estremo rigore sulle osservazioni preliminari, il dato sintetico rimane comunque una semplificazione della realtà.
- **Omogeneità della popolazione:** I soggetti osservati presentavano un profilo prevalentemente iper-reattivo. Di conseguenza, il modello attuale non è direttamente generalizzabile a profili ASD ipo-reattivi o con pattern autonomici atipici, che richiederebbero una ricalibrazione specifica.
- **Semplificazione della struttura dei segnali:** I dati sintetici, per quanto "sporcati" con rumore, non includono la totalità degli artefatti da movimento o le non-stazionarietà tipiche delle registrazioni reali prolungate. Le performance riportate (AUC, sensibilità) devono quindi essere interpretate come stime ottimali in condizioni controllate.
- **Approssimazione del label shifting:** L'uso di un intervallo fisso per anticipare l'allarme è una scelta progettuale per garantire l'anticipo, ma nella realtà clinica il tempo di accumulo dello stress è variabile. Per affinare questa precisione, il sistema dovrà auto-aggiornarsi nel tempo, ricalibrando continuamente le proprie soglie in base all'evoluzione fisiologica del singolo paziente nel corso dei giorni e dei mesi.

11.3 Sviluppi futuri: verso il transfer learning

Le limitazioni riscontrate indicano chiaramente la strada per il futuro del progetto. L'obiettivo primario è sicuramente la raccolta di un dataset clinico più vasto, che includa almeno 10-15 partecipanti monitorati durante intere sessioni di vita quotidiana.

Tuttavia, la vera svolta metodologica per l'analisi di questi futuri dati potrebbe consistere nell'implementazione del *Transfer Learning* (apprendimento per trasferimento). In ambito di *Machine Learning*, questa tecnica permette di non dover addestrare un algoritmo partendo da zero ogni volta che cambia il contesto operativo, ma di "trasferire" la conoscenza già acquisita in un dominio di partenza a un dominio nuovo e più complesso. In quest'ottica, il modello sviluppato in questa tesi non verrà scartato o sostituito, ma fungerà da solida struttura pre-addestrata. I dati sintetici gli hanno infatti già fornito le "fondamenta": il sistema ha imparato le regole biologiche e matematiche universali dello stress (come la correlazione tra l'aumento della pendenza EDA e il ritiro vagale). I futuri dati clinici reali, che sono più difficili da raccogliere e molto più "sporchi", non dovranno quindi insegnare all'algoritmo cos'è una crisi partendo dal nulla. Essi, infatti, serviranno a specializzare l'algoritmo, abituardolo a riconoscere quelle stesse regole biologiche in mezzo al caos della quotidianità e insegnandogli a non farsi ingannare dai classici rumori di fondo della vita reale (come lo sfregamento del sensore, le variazioni termiche o i movimenti casuali). In sintesi, il presente lavoro ha permesso di strutturare e validare teoricamente un'architettura predittiva solida e di dimostrare che la fisiologia dello stress segue regole universali, decodificabili sia nel *cockpit* di un aereo sia nelle sfide quotidiane della neurodivergenza. Il naturale passo successivo consisterà nella sua calibrazione diretta sul campo. L'auspicio è che questo passaggio possa tradurre l'impianto algoritmico sviluppato in uno strumento di supporto efficace, contribuendo concretamente a migliorare la sicurezza, l'autonomia e la quotidianità delle persone con autismo e delle loro famiglie.

Bibliografia

- [1] Matthew S Goodwin et al. “Predicting imminent aggression onset in minimally-verbal youth with autism spectrum disorder using preceding physiological signals”. In: *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 2018, pp. 201–207.
- [2] Matthew S Goodwin et al. “Predicting aggression to others in youth with autism using a wearable biosensor”. In: *Autism research* 12.8 (2019), pp. 1286–1296.
- [3] Tales Imbiriba et al. “Wearable biosensing to predict imminent aggressive behavior in psychiatric inpatient youths with autism”. In: *JAMA network open* 6.12 (2023), e2348898.
- [4] Stephen M Kanne e Micah O Mazurek. “Aggression in children and adolescents with ASD: Prevalence and risk factors”. In: *Journal of autism and developmental disorders* 41.7 (2011), pp. 926–937.
- [5] Karla Conn Welch. “Physiological signals of autistic children can be useful”. In: *IEEE Instrumentation & Measurement Magazine* 15.1 (2012), pp. 28–32.
- [6] Matthew S Goodwin et al. “Cardiovascular arousal in individuals with autism”. In: *Focus on autism and other developmental disabilities* 21.2 (2006), pp. 100–123.
- [7] Bruce S McEwen. “Protective and damaging effects of stress mediators”. In: *New England journal of medicine* 338.3 (1998), pp. 171–179.
- [8] Azadeh Kushki et al. “Investigating the autonomic nervous system response to anxiety in children with autism spectrum disorders”. In: *PLoS one* 8.4 (2013), e59730.
- [9] Bradley J Ferguson et al. “Examining the association between electrodermal activity and problem behavior in severe autism spectrum disorder: A feasibility study”. In: *Frontiers in psychiatry* 10 (2019), p. 654.

- [10] Jessica L Greenlee et al. “Comparative analysis of electrodermal activity metrics and their association with child behavior in autism spectrum disorder”. In: *Developmental psychobiology* 66.2 (2024), e22461.
- [11] Sakeena Panju et al. “Atypical sympathetic arousal in children with autism spectrum disorder and its association with anxiety symptomatology”. In: *Molecular autism* 6.1 (2015), p. 64.
- [12] Emily Barbara Prince et al. “The relationship between autism symptoms and arousal level in toddlers with autism spectrum disorder, as measured by electrodermal activity”. In: *Autism* 21.4 (2017), pp. 504–508.
- [13] Angelina Verneti et al. “Atypical emotional electrodermal activity in toddlers with autism spectrum disorder”. In: *Autism Research* 13.9 (2020), pp. 1476–1488.
- [14] June Groden et al. “Assessing cardiovascular responses to stressors in individuals with autism spectrum disorders”. In: *Focus on Autism and Other Developmental Disabilities* 20.4 (2005), pp. 244–252.
- [15] Rachel M Fenning et al. “Sympathetic-parasympathetic interaction and externalizing problems in children with autism spectrum disorder”. In: *Autism Research* 12.12 (2019), pp. 1805–1816.
- [16] Azadeh Kushki et al. “A Kalman filtering framework for physiological detection of anxiety-related arousal in children with autism spectrum disorder”. In: *IEEE Transactions on Biomedical Engineering* 62.3 (2014), pp. 990–1000.
- [17] David Rosenfield et al. “Change point analysis for longitudinal physiological data: Detection of cardio-respiratory changes preceding panic attacks”. In: *Biological psychology* 84.1 (2010), pp. 112–120.
- [18] Patient Zihisire Muke e Adrianna Koziarkiewicz. “Machine learning techniques to improve the cognitive workload classification using multimodal sensors’ data”. In: *IEEE Access* (2025).
- [19] Elena-Simona Apostol et al. “Change point enhanced anomaly detection for IoT time series data”. In: *Water* 13.12 (2021), p. 1633.

- [20] Hassan Ismail Fawaz et al. “Data augmentation using synthetic data for time series classification with deep residual networks”. In: *arXiv preprint arXiv:1808.02455* (2018).
- [21] Lucas Lange, Nils Wenzlitschke e Erhard Rahm. “Generating synthetic health sensor data for privacy-preserving wearable stress detection”. In: *Sensors* 24.10 (2024), p. 3052.
- [22] Farida Sabry et al. “Wearable Data Generation Using Time-Series Generative Adversarial Networks for Hydration Monitoring.” In: *BIOSIGNALS*. 2023, pp. 94–105.
- [23] Md Abid Hasan et al. “A comprehensive survey and comparative analysis of time series data augmentation in medical wearable computing”. In: *Plos one* 20.3 (2025), e0315343.
- [24] Mathias Benedek e Christian Kaernbach. “A continuous measure of phasic electrodermal activity”. In: *Journal of neuroscience methods* 190.1 (2010), pp. 80–91.
- [25] Alberto Greco et al. “Electrodermal activity processing: A convex optimization approach”. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2014, pp. 2290–2293.
- [26] Carolyn McCormick et al. “Electrodermal and behavioral responses of children with autism spectrum disorders to sensory and repetitive stimuli”. In: *Autism Research* 7.4 (2014), pp. 468–480.
- [27] Sinéad Lydon, Olive Healy e Martina Dwyer. “An examination of heart rate during challenging behavior in autism spectrum disorder”. In: *Journal of Developmental and Physical Disabilities* 25.1 (2013), pp. 149–170.
- [28] James A Russell. “A circumplex model of affect”. In: *Journal of personality and social psychology* 39.6 (1980), pp. 1161–1178.
- [29] Wolfram Boucsein. *Electrodermal activity*. Springer science & business media, 2012.

- [30] Michael E Dawson, Anne M Schell, Diane L Filion et al. “The electrodermal system”. In: *Handbook of psychophysiology* 2 (2007), pp. 200–223.
- [31] Fred Shaffer e Jay P Ginsberg. “An overview of heart rate variability metrics and norms”. In: *Frontiers in public health* 5 (2017), p. 290215.
- [32] SW Porges. *The polyvagal perspective Biological Psychology*, 74 (2), 116–143. 2007.
- [33] Varun Chandola, Arindam Banerjee e Vipin Kumar. “Anomaly detection: A survey”. In: *ACM computing surveys (CSUR)* 41.3 (2009), pp. 1–58.
- [34] Douglas C Montgomery. *Introduction to statistical quality control*. John wiley & sons, 2020.
- [35] Willis J Tompkins. *Biomedical digital signal processing*. Vol. 237. Prentice Hall New Jersey, 1993.
- [36] Leslie G Portney. *Foundations of clinical research: applications to evidence-based practice*. FA Davis, 2020.
- [37] Ming-Zher Poh, Nicholas C Swenson e Rosalind W Picard. “A wearable sensor for unobtrusive, long-term assessment of electrodermal activity”. In: *IEEE transactions on Biomedical engineering* 57.5 (2010), pp. 1243–1252.
- [38] Leo Breiman. “Random forests”. In: *Machine learning* 45.1 (2001), pp. 5–32.
- [39] V Rajendra Mali e Anil R Surve. “Stress prediction using random forest linear regression”. In: *Int. J. Eng. Appl. Sci. Technol.* 6.3 (2021), pp. 1–15.
- [40] Stephanie Hödl et al. “Pre-ictal heart rate variability alterations in focal onset seizures and response to vagus nerve stimulation”. In: *Seizure* 86 (2021), pp. 175–180.
- [41] Stefania Coelli et al. “Modulation of autonomic responses to cognitive tasks under acute mental stress”. In: *Scientific Reports* (2026).
- [42] Kayisan M Dalmeida e Giovanni L Masala. “HRV features as viable physiological markers for stress detection using wearable devices”. In: *Sensors* 21.8 (2021), p. 2873.

- [43] Shaked Rosenblum, Sharona L Rab e Roee Admon. “Dynamics in physiological acute stress response trajectories: uncovering latent variability”. In: *BMC psychiatry* 25.1 (2025), p. 361.
- [44] Peizheng Wang, Robert Houghton e Arnab Majumdar. “Detecting and predicting pilot mental workload using heart rate variability: a systematic review”. In: *Sensors* 24.12 (2024), p. 3723.
- [45] Xiaoru Yu et al. “Exploring physiological stress response evoked by passive translational acceleration in healthy adults: a pilot study utilizing electrodermal activity and heart rate variability measurements”. In: *Scientific reports* 14.1 (2024), p. 11349.
- [46] Empatica S.r.l. *Empatica E4 Technical Specifications (Revision 001)*. Technical datasheet. Empatica S.r.l. Milano, Italy, nov. 2014. URL: <https://box.empatica.com>.
- [47] Empatica Inc. *Sito ufficiale e documentazione tecnica*. <https://www.empatica.com>. 2026.
- [48] Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. “Heart rate variability: standards of measurement, physiological interpretation, and clinical use”. In: *Circulation* 93.5 (1996), pp. 1043–1065.
- [49] Maurizio Garbarino et al. “Empatica E3—A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition”. In: *2014 4th international conference on wireless mobile communication and healthcare-transforming healthcare through innovations in mobile and wireless technologies (MOBIHEALTH)*. IEEE. 2014, pp. 39–42.
- [50] Empatica Inc. *Empatica Health Monitoring Platform - Data Access Guide*. PRS-0000205 REV27.0. Empatica Inc. Boston, MA, 2025. URL: <https://www.empatica.com>.

Ringraziamenti