



UNIVERSITÀ
DI PAVIA

Dipartimento di Biologia e Biotecnologie "L. Spallanzani"

Laurea Magistrale in Biotecnologie Avanzate Plus

Azienda "Vegea Srl"

Progettazione di un dispositivo per il monitoraggio
del deperimento di scarti vitivinicoli e valutazione
preliminare mediante modelli predittivi

Relatore Interno:
Professoressa Monica Savio

Tutor aziendale:
Dott. Francesco Merlino

Tesi Sperimentale di
Annalisa Simone

Anno Accademico 2024/2025

Riassunto

L'industria vitivinicola rappresenta uno dei principali settori agroalimentari a livello globale. Secondo il report 2025 dell'Organizzazione Internazionale della Vigna e del Vino, riferito allo scenario del 2024, la produzione mondiale di vino è stimata in circa 226 milioni di ettolitri. Durante il processo di vinificazione vengono generati tra 1,3 e 1,5 kg di rifiuti per litro di vino prodotto; tra i residui organici, la vinaccia costituisce il principale sottoprodotto, risultando ricca di composti di interesse ma altamente deperibile. La sua corretta caratterizzazione e selezione risulta pertanto cruciale per una valorizzazione sostenibile ed efficiente.

Il presente lavoro si inserisce in un progetto volto allo sviluppo di un dispositivo per il monitoraggio della qualità della vinaccia. A tal fine, sono state selezionate le principali variabili da monitorare, tra cui temperatura, umidità relativa, pH, tempo e intensità luminosa. In un'ottica di futura implementazione, sono stati testati diversi modelli di intelligenza artificiale per individuare l'approccio più idoneo alla previsione della *shelf life* della vinaccia sulla base dei parametri ambientali misurati dal dispositivo.

Sono stati valutati modelli di regressione tradizionali (*Linear Regression*, *Random Forest*, *Gradient Boosting* e *Support Vector Regression*) e modelli basati su reti neurali (*Multilayer Perceptron Regressor* e *ResNet* implementate in Keras), utilizzando un dataset fittizio costruito sulla base dei sensori del dispositivo, al fine di condurre un'analisi preliminare e individuare il modello di intelligenza artificiale più idoneo alla futura integrazione nel sistema. I risultati hanno evidenziato che la rete neurale *Multilayer Perceptron Regressor* implementata in Keras, con due strati nascosti, ha fornito le migliori prestazioni predittive complessive ($R^2 = 0,864$, MAE = 2,33, MSE = 9,22).

Successivamente, è stato introdotto il contenuto di antocianine come parametro aggiuntivo di qualità, ipotizzato come indicatore chimico del degrado della vinaccia nel tempo. L'analisi esplorativa ha mostrato una forte correlazione positiva tra il livello di antocianine e la *shelf life* ($r = 0,85$). L'inclusione di tale variabile nel dataset fittizio ha determinato un miglioramento coerente delle prestazioni predittive dei modelli; in particolare, la rete neurale MLP in Keras ha raggiunto valori pari a $R^2 = 0,890$, MAE = 1,23 e MSE = 3,81. L'analisi bootstrap ha evidenziato una tendenza sistematica al miglioramento delle prestazioni in presenza della variabile "antocianine", sebbene con livelli di significatività statistica prossimi alla soglia convenzionale (p compreso tra 0,06 e 0,12).

Nel complesso, i risultati suggeriscono che l'integrazione di modelli di intelligenza artificiale e di parametri chimici aggiuntivi, come il contenuto di antocianine, possa migliorare l'efficacia

dei sistemi di monitoraggio della vinaccia, fornendo una base solida per lo sviluppo di tecnologie dedicate alla valorizzazione sostenibile dei sottoprodotti vitivinicoli.

INDICE

1. INTRODUZIONE	1
1.1 CONTESTO DELL'INDUSTRIA VITIVINICOLA E SOSTENIBILITÀ	1
1.2 STRATEGIE DI VALORIZZAZIONE DEGLI SCARTI VITIVINICOLI	3
1.3 COMPOSTI BIOATTIVI NEGLI SCARTI VITIVINICOLI	5
1.4 APPLICAZIONI INDUSTRIALI DEI COMPOSTI BIOATTIVI DELLA VINACCIA	7
1.5 SHELF LIFE DEGLI SCARTI VITIVINICOLI E POTENZIALI INDICATORI	9
1.5.1 COMPOSTI FENOLICI E VARIAZIONI CROMATICHE	10
1.5.2 LA TEMPERATURA	11
1.5.3 UMIDITÀ RELATIVA	11
1.5.4 VOCs	12
1.5.5 CONTAMINAZIONI MICROBIOLOGICHE E MICOTOSSINE	13
1.6 EVOLUZIONE DEI METODI PREDITTIVI DELLA SHELF LIFE DI BIOMASSA DEPERIBILE	14
1.7 MODELLI DI INTELLIGENZA ARTIFICIALE PER LA SHELF LIFE DI BIOMASSA VITIVINICOLA	15
1.8 MODELLI DI INTELLIGENZA ARTIFICIALE UTILIZZATI IN QUESTO STUDIO	19
1.8.1 REGRESSIONE LINEARE MULTIPLA (MLR)	19
1.8.2 RANDOM FOREST REGRESSOR (RFR)	20
1.8.3 GRADIENT BOOSTING REGRESSOR (GBR)	21
1.8.4 SUPPORT VECTOR REGRESSION (SVR)	22
1.8.5 RETI NEURALI ARTIFICIALI (MULTILAYER PERCEPTRON, MLP) .	23
1.8.6 RETI NEURALI DEEP LEARNING (RESIDUAL NETWORK, RESNET)	24
1.9 BOOTSTRAPPING	26
1.10 SENSORI E IOT	27
2. SCOPO DEL LAVORO	30
3. MATERIALI E METODI	31
3.2 COMPONENTE HARDWARE	31
3.2.1 MICROCONTROLLORE PRINCIPALE - ESP32-S3-WROOM-1	31
3.1.2 SENSORE DI TEMPERATURA E UMIDITÀ - SHT31	32
3.1.3 REAL TIME CLOCK - DS3231	32
3.1.4 SENSORE DI LUMINOSITÀ – TSL2591	32
3.1.5 SENSORE DI pH	33
3.1.6 MICROSD - MEMORIA ESTERNA	34
3.1.7 MODULO DI RICARICA BATTERIA - TP4056	34
3.1.8 MODULO MONITORAGGIO BATTERIA - MAX 17043	35
3.1.9 FIRMWARE DI CONTROLLO DEL DISPOSITIVO DI MONITORAGGIO	35
3.1.10 COMUNICAZIONE ARDUINO E SOFTWARE DI SUPERVISIONE ...	36
3.2 COMPONENTE SOFTWARE	38

3.2.1	BACKEND	38
3.2.2	FRONTEND	40
3.2.4	GESTIONE E RICEZIONE DEI DATI NEL SOFTWARE DI SUPERVISIONE	41
3.3	MODULO DI INTELLIGENZA ARTIFICIALE	42
3.3.1	AMBIENTE DI SVILUPPO E STRUMENTI SOFTWARE.....	42
3.3.2	GENERAZIONE DEI DATASET SINTETICI	43
3.3.3	ANALISI ESPLORATIVA DEI DATI SUL LIVELLO DI ANTOCIANINE	44
3.3.4	MODELLI DI INTELLIGENZA ARTIFICIALE UTILIZZATI.....	45
3.3.4.1	REGRESSIONE LINEARE MULTIPLA	45
3.3.4.2	<i>RANDOM FOREST REGRESSOR (RFR)</i>	45
3.3.4.3	<i>GRADIENT BOOSTING REGRESSOR (GBR)</i>	45
3.3.4.4	<i>SUPPORT VECTOR REGRESSION (SVR)</i>	45
3.3.4.5	<i>MULTILAYER PERCEPTRON REGRESSOR IN SCIKIT-LEARN</i>	46
3.3.4.6	<i>MULTILAYER PERCEPTRON REGRESSOR IN KERAS</i>	46
3.3.4.7	<i>IMPROVED RESNET IN KERAS</i>	46
3.3.5	METRICHE DI VALUTAZIONE DELLE PRESTAZIONI DEI MODELLI	47
3.3.6	<i>BOOTSTRAPPING</i>	48
4.	RISULTATI E DISCUSSIONE	50
4.1	DISPOSITIVO IOT SVILUPPATO.....	50
4.2	RISULTATI ANALISI COMPARATIVA DEI MODELLI DI IA TESTATI (<i>DATASET A</i>).....	51
4.3	ANALISI ESPLORATIVA SULLA VARIABILE “ANTOCIANINE”	54
4.3.1	STATISTICHE DESCRITTIVE E DISTRIBUZIONE	55
4.3.2	INTERPRETAZIONE DELLE CORRELAZIONI	56
4.4	INTEGRAZIONE DEL MODULO ANTOCIANINE E NUOVI RISULTATI (<i>DATASET B</i>).....	59
4.5	ANALISI BOOTSTRAP COMPARATIVA.....	62
5.	CONCLUSIONI	69
6.	BIBLIOGRAFIA	72

Abbreviazioni

	Inglese	Italiano
3D-CNN	3D Convolutional Neural Network	Rete Neurale Convoluzionale 3D
ASLT	Accelerated Shelf Life Test	Test Accelerato di Shelf Life
BLE	Bluetooth Low Energy	Bluetooth a basso consumo energetico
CO ₂	Carbon dioxide	Anidride Carbonica
DAE	Denoising Autoencoder	Autoencoder per rimozione del rumore
DCAE	Deep Convolutional Autoencoder	Autoencoder Convoluzionale Profondo
DHS	4,4'-dihydroxy-trans-stilbene	4,4'-diidrossi-trans-stilbene
DL	Deep Learning	Apprendimento Profondo
EDA	Exploratory Data Analysis	Analisi Esplorativa dei Dati
GBR	Gradient Boosting Regressor	Regressore Gradient Boosting
IA	Artificial Intelligence	Intelligenza Artificiale
IoT	Internet of Things	Internet delle Cose
LDL	Low-Density Lipoprotein	Lipoproteine a Bassa Densità
LPP	Locality Preserving Projection	Proiezione a Conservazione Locale
LR	Linear Regression	Regressione Lineare
LSTM	Long Short-Term Memory	Memoria a Lungo Termine

LLC	Lewis Lung Carcinoma	Carcinoma Polmonare Lewis
MAE	Mean Absolute Error	Errore Assoluto Medio
MASLT	Multivariate Accelerated Shelf Life Test	Test Accelerato Multivariato di Shelf Life
ML	Machine Learning	Apprendimento Automatico
MLP	Multilayer Perceptron	Perceptrone Multistrato
MLPRegressor	Multilayer Perceptron Regressor	Perceptrone Multistrato regressore
MSE	Mean Squared Error	Errore Quadratico Medio
NFC	Near Field Communication	Comunicazione in prossimità
OTA	Ochratoxin A	Ocratossina A
PCA	Principal Component Analysis	Analisi delle Componenti Principali
PLS	Partial Least Squares	Regressione a Minimi Quadrati Parziali
PUFA	Polyunsaturated Fatty Acids	Acidi Grassi Polinsaturi
RBF	Radial Basis Function	Funzione a Base Radiale
ResNET	Residual Network	Rete Neurale Residuale
RF	Random Forest	Foresta Casuale
RFID	Radio Frequency Identification	Identificazione a Radiofrequenza
RFR	Random Forest Regressor	Regressore Random Forest

SDAE	Stacked Denoising Autoencoder	Autoencoder Denoising Impilato
SoC	State of Charge	Stato di Carica
SVR	Support Vector Regressor	Regressore a Vettori di Supporto
UV	Ultraviolet	Ultravioletto
Vis/NIR	Visible / Near-Infrared	Spettro Visibile / Vicino Infrarosso
VOCs	Volatile Organic Compounds	Composti Organici Volatili
VNIR-SWIR	Visible–Near Infrared / Short-Wave Infrared	Spettro Visibile–Vicino Infrarosso / Infrarosso a Onde Corte
VPD	Vapor Pressure Deficit	Deficit di Pressione di Vapore

1. INTRODUZIONE

1.1 CONTESTO DELL'INDUSTRIA VITIVINICOLA E SOSTENIBILITÀ

L'industria di produzione del vino è tra le principali industrie nel settore agro-alimentare che genera guadagni significativi su scala globale. (Laca *et al.*, 2021) Secondo il report del 2025 dell'Organizzazione Internazionale della Vigna e del Vino, incentrato sul precedente scenario del 2024, la produzione di vino mondiale nel 2024 è stimata essere 226 milioni di ettolitri. (Report 2025) Nello specifico, l'Italia (44.1 milioni di ettolitri), Francia (36.1 milioni di ettolitri) e Spagna (31.0 milioni di ettolitri), rappresentano il 50% della produzione mondiale nel 2024, una quota che da sola si avvicina a quella dell'intera Europa, pari al 61%. (Report 2025)

In molte regioni del mondo, la viticoltura e la produzione di vino rappresentano un importante patrimonio culturale sia a livello economico che sociale. (Santillán *et al.*, 2019) Tuttavia, la viticoltura e la produzione di vino influenzano l'ambiente come qualsiasi altra attività economica. (Viveros Santos *et al.*, 2023) Per questo motivo l'industria del vino è uno dei sistemi agro-alimentari più studiati da un punto di vista della sostenibilità ambientale, data la sua importanza alla produzione economica primaria e il suo commercio internazionale. (Abbate *et al.*, 2025)

Il settore vitivinicolo è fortemente associato alla sostenibilità ambientale perché arricchisce il suolo che supporta la vita e la crescita di una vasta gamma di piante e per la sua suscettibilità ai cambi di temperatura. (Ferrer *et al.* 2022) La produzione di vino ha anche numerosi effetti avversi sull'ambiente, come la produzione di scarti e lo spreco di acqua, l'utilizzo di pesticidi e di terreni, impiego di energia, così come la *carbon footprint* del trasporto. (Brito *et al.*, 2024) Per questo motivo, la filiera del vino sta gradualmente passando a metodi di produzione più sostenibili. (Casson *et al.*, 2022) In questo contesto, integrando considerazioni economiche, ambientali e sociali, le pratiche di economia circolare offrono un percorso verso un futuro più sostenibile e resiliente, allineandosi con l'agenda 2030 globale dello Sviluppo Sostenibile delle Nazioni Unite. (Matrapazi *et al.*, 2020)

Nonostante la crescente attenzione verso la sostenibilità nell'industria vitivinicola, permane la necessità di un'applicazione più concreta di queste tecniche, poiché diversi fattori contribuiscono al divario tra teoria e pratica, come la difficoltà di bilanciare le dimensioni ambientali, sociali ed economiche della sostenibilità e la necessità di colmare le lacune di

conoscenza sui benefici e sui costi percepiti in termini di tempo, risorse e tecnologia. (Baiano 2021) Alcuni produttori di vino, in particolare le realtà più piccole, potrebbero incontrare difficoltà nell'adottare pratiche sostenibili a causa di vincoli finanziari, della mancanza di conoscenze sulle soluzioni disponibili e della limitata partecipazione a reti e *partnership* locali che facilitano la diffusione di approcci sostenibili. (Casini *et al.*, 2010) In secondo luogo, coordinare pratiche sostenibili lungo l'intera catena di produzione può essere una sfida, poiché richiede collaborazione, comunicazione e allineamento di obiettivi e priorità. (Balaine *et al.*, 2023)

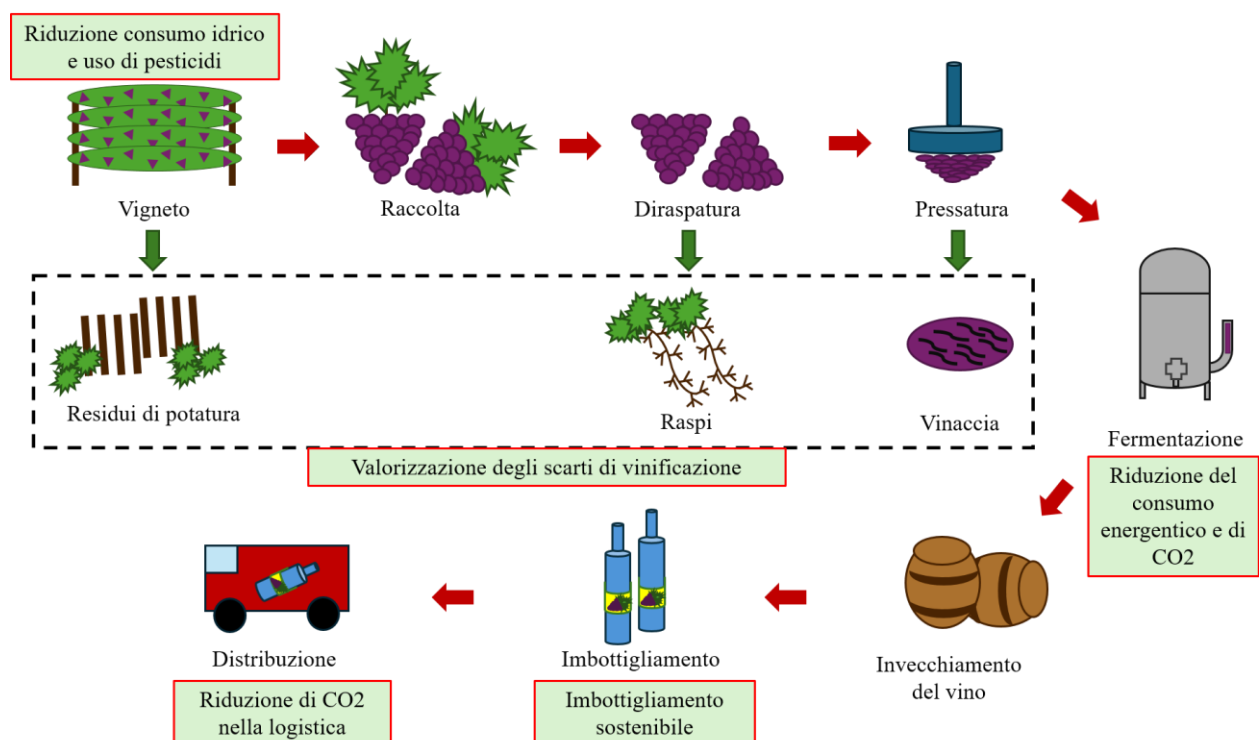


Figura 1.1 Schema del processo produttivo del vino e dei principali punti di intervento per migliorare la sostenibilità lungo la filiera, dalla gestione del vigneto alla distribuzione finale (riquadri verdi). (Adattato da Jesus *et al.*, 2022)

Si può rendere più sostenibile la filiera vitivinicola intervenendo in diversi punti del processo di produzione, dalla produzione vera e propria, alla commercializzazione e all'imballaggio sostenibili, come riportato in **Figura 1.1**. (Maesano *et al.*, 2022) In più, la catena di produzione del vino dovrebbe includere anche processi sostenibili per lo smaltimento dei rifiuti, il riutilizzo dei sottoprodotti e la loro trasformazione per generare altri beni e/o energia basandosi sull'economia circolare, un *framework* economico progettato per autorigenerarsi. (Ezeudu *et al.*, 2021) In questo contesto si inserisce il concetto di "cantina rigenerativa", che va oltre la

sostenibilità tradizionale per abbracciare un approccio capace di ripristinare gli ecosistemi e chiudere i cicli delle risorse. (Costa *et al.*, 2022)

Adottando una strategia economica di tipo circolare, in cui tutti i rifiuti di produzione vengono trasformati in una risorsa preziosa e materie specifiche possono essere trasferite da un sistema di produzione a un altro senza essere smaltite, è possibile ridurre l'impatto ambientale del processo di vinificazione. (Perra *et al.*, 2022)

1.2 STRATEGIE DI VALORIZZAZIONE DEGLI SCARTI VITIVINICOLI

Durante il processo di produzione del vino vengono generati tra 1,3 e 1,5 kg di rifiuti per litro di vino prodotto, di cui il 75% è acqua reflua. (Lucas *et al.*, 2010) Gli altri principali residui derivanti dall'attività di vinificazione sono rappresentati da: rifiuti organici (bucce di uva, che contengono semi, polpa e bucce, raspi e foglie di uva), fecce di vino, emissioni di gas serra (CO₂, composti organici volatili, ecc.), e rifiuti inorganici (terra di diatomee, argilla bentonitica e perlite). (Maicas *et al.*, 2020) In totale, ogni anno in Europa vengono prodotti circa 14,5 milioni di tonnellate di sottoprodotti dell'uva. (Chouchou *et al.*, 2013)

Considerata l'elevata quantità di residui prodotti dalla vinificazione e il loro impatto ambientale, negli ultimi anni sono state sviluppate numerose strategie volte alla loro valorizzazione, trasformando questi sottoprodotti in risorse utili per diversi settori industriali e contribuendo alla sostenibilità del processo produttivo. (Muhlack *et al.*, 2018)

La **Figura 1.2** schematizza i principali flussi di scarto della vinificazione e le possibili modalità di riutilizzo e recupero orientate alla sostenibilità del processo produttivo.

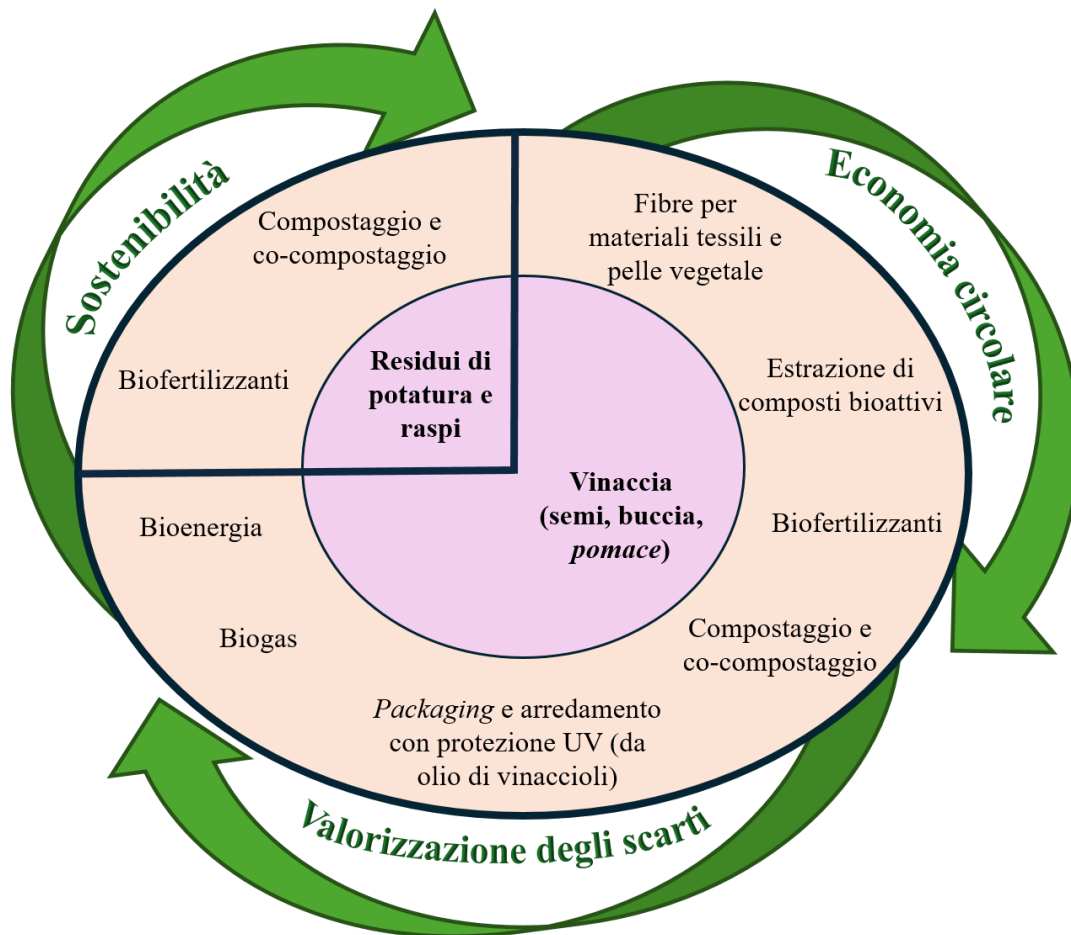


Figura 1.2 Sottoprodotti della vinificazione e strategie di valorizzazione secondo l'economia circolare.

Esistono tante modalità per valorizzare gli scarti dell'industria vitivinicola, come la produzione di bioenergia, che consente di ridurre la dipendenza dai combustibili fossili non rinnovabili e di abbattere le emissioni di gas serra. (Kassongo *et al.*, 2020) Un altro esempio è la produzione di biogas ottenuto per fermentazione anaerobica della materia organica presente nei residui vitivinicoli. (Donia *et al.*, 2018)

Il compostaggio o il co-compostaggio di vinaccia, raspi e fanghi provenienti dagli impianti di trattamento delle acque reflue fornisce un'alternativa ecologica, utile sia per migliorare la fertilità del suolo sia come soluzione sostenibile per lo smaltimento finale dei rifiuti vinicoli. (Gómez-Brandón *et al.*, 2021)

Un'ulteriore opzione per il riciclo di questi residui, in particolare delle vinacce e dei raspi d'uva, consiste nel loro impiego come fertilizzanti organici, sia mediante applicazione diretta al suolo sia dopo compostaggio, oppure come substrati per la coltivazione. (Bustamante *et al.*, 2008)

L'olio di vinaccioli, che è stato dimostrato conferire ai film d'imballaggio una colorazione opaca e una potenziale protezione degli alimenti dalla degradazione causata dai raggi UV, può essere estratto e impiegato come materia prima per la realizzazione di nuovi materiali destinati

all'arredamento o al *packaging*, sebbene siano necessari ulteriori approfondimenti sperimentali per confermarne pienamente le potenzialità. (Mauro *et al.*, 2022)

Le fibre presenti nell'uva, che possono rappresentare una risorsa preziosa nel settore della moda come alternativa ecologica e durevole alla pelle tradizionale, sono state oggetto degli studi dell'architetto Gianpiero Tessitore, il quale ha identificato quelle più adatte alla realizzazione di una pelle vegetale ed ecosostenibile, dando così origine al progetto "WineLeather". (Abbate *et al.*, 2025) Alla base dell'iniziativa vi è l'azienda Vegea Srl, azienda ospitante della presente esperienza curriculare e che ha saputo trasformare i sottoprodotti della vinificazione in una materia innovativa destinata all'industria tessile e della moda sostenibile. (VEGEA S.r.l., n.d.) Tra le strategie più promettenti rientra l'estrazione dei composti bioattivi dagli scarti vitivinicoli, un approccio che ha attirato l'interesse di numerose aziende enologiche per le sue potenzialità applicative in diversi settori industriali. (Abreu *et al.*, 2024)

1.3 COMPOSTI BIOATTIVI NEGLI SCARTI VITIVINICOLI

I composti bioattivi sono sostanze presenti in natura, fanno parte della catena alimentare e si trovano generalmente in piccole quantità negli alimenti, ma sono in grado di esercitare effetti benefici sulla salute umana grazie alla loro influenza su attività fisiologiche e cellulari. (Biesalski *et al.*, 2009)

Tra i frutti con una concentrazione particolarmente elevata di composti bioattivi vi è l'uva, i cui benefici per la salute sono principalmente associati alla presenza di composti fenolici, noti per l'elevato potenziale antiossidante, e di acidi grassi polinsaturi (PUFA), in particolare l'acido linoleico e l'acido α -linolenico, presenti soprattutto nei semi. (Riberiro *et al.*, 2015)

I composti fenolici si suddividono in due grandi gruppi: i non flavonoidi (come acidi idrossibenzoici, idrossicinnamici e stilbeni) e i flavonoidi (tra cui antociani, flavonoli, flavan-3-oli liberi o polimerizzati e tannini). (Montealegre *et al.*, 2006)

Per quanto riguarda i grassi, i principali PUFA sono appunto l'acido linoleico e l'acido α -linolenico, considerati essenziali poiché coinvolti in importanti attività energetiche, metaboliche e strutturali. (Lobb *et al.*, 2007)

Nell'uva, i composti fenolici risultano distribuiti per circa il 10% nella polpa, il 60–70% nei semi e il 28–35% nelle bucce. Tuttavia, la loro concentrazione può variare significativamente in base alla varietà genetica dell'uva e all'influenza del processo di vinificazione. (Riberiro *et al.*, 2015)

La **Figura 1.3** illustra la distribuzione dei principali composti chimici presenti nell'uva rossa.

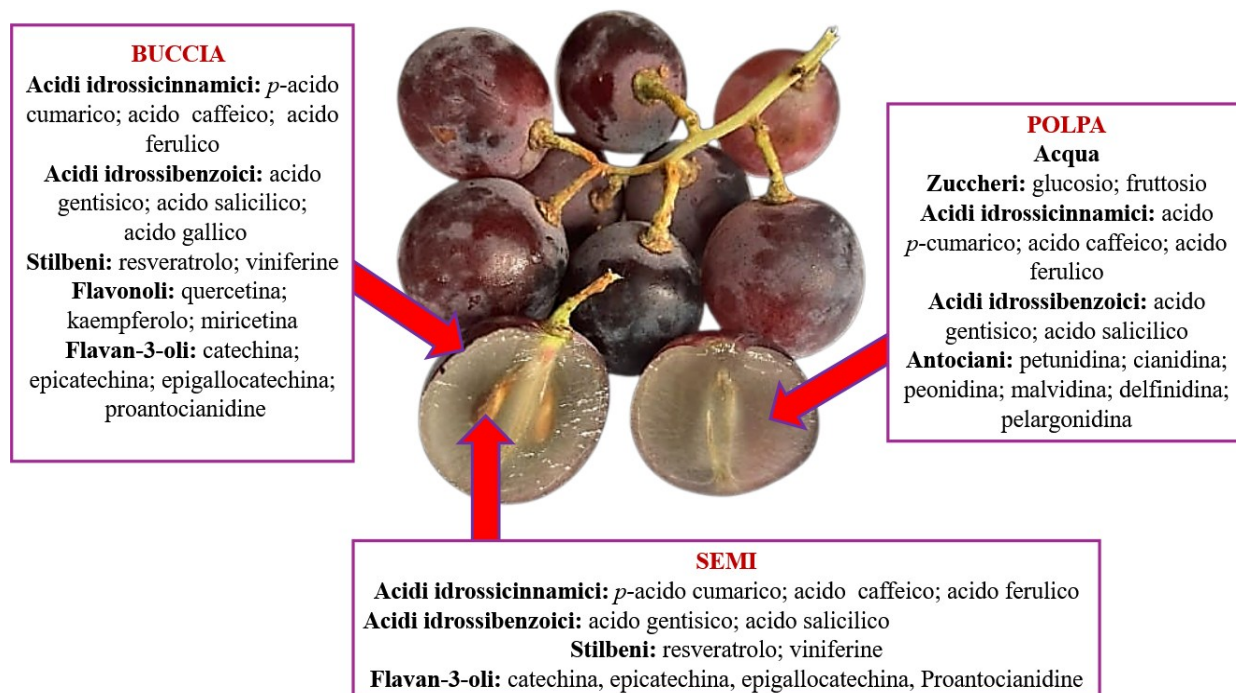


Figura 1.3. Distribuzione dei principali composti chimici nell'uva rossa. (Riprodotta da Nirmal *et al.*, 2025)

Un aspetto particolarmente rilevante è che i composti fenolici, di grande valore per le loro proprietà bioattive, rimangono presenti in quantità significative anche dopo la vinificazione, a causa dell'estrazione incompleta che avviene durante il processo enologico (Castellanos-Gallo *et al.*, 2022)

Come mostra la Tabella I, i sottoprodotti che costituiscono la vinaccia [bucce (~50%), semi (~25%) e raspi (~25%)] non sono semplici scarti, in quanto sono estremamente ricchi di composti bioattivi, come polifenoli, flavonoidi, fibre alimentari e acidi organici, che ne fanno un'interessante risorsa per applicazioni nutraceutiche, cosmetiche, farmaceutiche e ambientali. (Karastergiou *et al.*, 2024)

Tabella I. Composizione della vinaccia. (Dati da Karastergiou *et al.*, 2024)

Componente	Quantità
Umidità (acqua)	50-60%
Fibra alimentare	20-30%
Carboidrati	5-15%
Proteine	8-15%
Polifenoli	5-10%
Lipidi	3-7%
Acidi organici	3-6%
Ceneri	3-5%

Secondo i dati presenti in letteratura, bucce e semi d'uva rappresentano una fonte rilevante di composti fenolici, con valori fino a 10.216 µg/g nei semi e una predominanza di antocianine nelle bucce, di cui solo una parte è trasferita nel vino finito. (Peixoto *et al.*, 2018)

Nonostante ciò, la distribuzione quantitativa e qualitativa dei composti fenolici nella vinaccia può variare significativamente in base a numerosi fattori: varietà dell'uva, condizioni climatiche, localizzazione geografica, tipo di suolo, influenze stagionali, grado di maturazione, tecniche viticole e modalità di vinificazione. (Kallithraka *et al.*, 2001)

Proprio la rilevanza e la variabilità del contenuto fenolico hanno spinto le ricerche più recenti a concentrarsi sulla valorizzazione della vinaccia come risorsa sostenibile. (Wang *et al.*, 2024)

1.4 APPLICAZIONI INDUSTRIALI DEI COMPOSTI BIOATTIVI DELLA VINACCIA

Grazie all'impiego di tecniche specifiche, oggi in continua evoluzione verso approcci più sostenibili, è possibile valorizzare la vinaccia e i semi d'uva attraverso l'estrazione di composti bioattivi di elevato interesse, quali oli, estratti fenolici e fibre alimentari, la cui applicazione trova spazio in diversi settori industriali. (Sirohi *et al.*, 2020) In particolare, i composti fenolici così estratti dalle vinacce, ammesso che abbiano preservato le loro attività biologiche, sono adatti per l'uso nelle industrie alimentari, cosmetiche e farmaceutiche. (Abbate *et al.*, 2025)

Infatti, i semi d'uva sgrassati e le bucce, ricche di polifenoli, rappresentano un'eccellente risorsa per la produzione di alimenti funzionali. (Wang *et al.*, 2024) Oltre a essere naturalmente privi di glutine, questi sottoprodotti sono anche una fonte preziosa di fibra alimentare. (Di Stefano *et*

al., 2023) Possono essere incorporati direttamente, senza necessità di lavorazioni preliminari, in prodotti alimentari di largo consumo come ingredienti bioattivi, valorizzando così gli scarti della filiera vitivinicola e permettendo ai produttori di ottenere un ritorno economico maggiore grazie alla percezione di questi ingredienti come prodotti ad alto valore aggiunto. (Abbate *et al.*, 2025)

Nell'ambito dell'industria cosmetica, farmaceutica e nutraceutica gli antociani, gli isoflavonoidi, i flavanoni, i flavanoli, gli acidi fenolici e i stilbeni presenti nell'uva sono noti per le loro proprietà farmacologiche, come antiossidanti, antimicrobici, anti-tirosinasi, vasodilatazione, antitumorale e antinfiammatoria. (Muhlack *et al.*, 2018) Queste proprietà bioattive degli estratti fenolici ottenuti dalla vinaccia hanno attirato l'attenzione della comunità scientifica per il trattamento o la prevenzione di specifiche patologie. (Machado *et al.*, 2024) Tra le più studiate vi sono le malattie cardiovascolari, poiché i composti fenolici contribuiscono al loro controllo riducendo i livelli di lipoproteine a bassa densità (LDL) e/o esercitando un'attività antiaggregante piastrinica. (Muñoz-Bernal *et al.*, 2021) I polifenoli presenti nella vinaccia (come tannini, flavonoidi, catechine e acido gallico) hanno dimostrato di possedere attività antiossidante e di poter inibire enzimi idrolitici come α -amilasi e α -glucosidasi, coinvolti nella digestione dei carboidrati, con possibili applicazioni nella gestione del diabete di tipo 2. (Kato-Schwartz *et al.*, 2020) Come riportato dallo studio di Lalegani *et al.* (2018), l'estratto idroalcolico di vinaccia di Merlot ha mostrato la capacità di inibire le α -amilasi salivare e pancreatica, nonché la α -glucosidasi intestinale, contribuendo così a ridurre l'assorbimento intestinale del glucosio e suggerendo potenziali applicazioni nella gestione del diabete di tipo 2. (Lalegani *et al.*, 2018)

Recentemente, nell'ambito delle malattie neurodegenerative, sono stati sviluppati nutriosomi, *nanocarrier* costituiti da Nutriose® (una maltodestrina ramificata solubile in acqua derivata dal mais) e fosfolipidi, per l'incapsulamento di estratti di vinaccia di uva Cannonau; questi sistemi hanno mostrato una maggiore efficacia nel proteggere le cellule epiteliali intestinali (Caco-2) dallo stress ossidativo indotto da perossido di idrogeno rispetto all'estratto libero, grazie alla superiore biodisponibilità dei polifenoli, aprendo nuove prospettive per l'impiego della vinaccia nel trattamento delle patologie neurodegenerative. (Allaw *et al.*, 2025)

Tra i composti bioattivi presenti nell'uva, spicca il resveratrolo, uno stilbene noto per le sue proprietà antiossidanti, antinfiammatorie e antitumorali, la cui efficacia è stata dimostrata da numerosi studi *in vitro* che ne hanno evidenziato la capacità di intervenire sia nelle fasi iniziali della trasformazione tumorale sia nella progressione del cancro, mostrando un ampio spettro d'azione. (Carter *et al.*, 2014) Il gruppo di Savio *et al.* ha ampiamente studiato il resveratrolo, valutandone anche l'efficacia rispetto a un analogo strutturale naturale, il 4,4'-dihydroxy-*trans*-

stilbene (DHS). E' stato dimostrato che la molecola è in grado di inibire *in vitro* la crescita delle cellule di carcinoma polmonare Lewis (LLC), inducendo apoptosi, arresto del ciclo cellulare in fase G1 e riducendo la migrazione e l'invasività delle cellule tumorali. (Savio *et al.*, 2016) In un successivo lavoro, lo stesso gruppo ha approfondito i meccanismi d'azione di DHS, evidenziandone il ruolo nella modulazione del microambiente tumorale, aprendo interessanti prospettive non solo sull'impiego del resveratrolo, ma anche sulla valutazione dei suoi analoghi, e ampliando così il potenziale applicativo dei composti bioattivi derivati dagli scarti della vinificazione. (Savio *et al.*, 2022)

1.5 SHELF LIFE DEGLI SCARTI VITIVINICOLI E POTENZIALI INDICATORI

La valorizzazione degli scarti vitivinicoli attraverso l'estrazione dei composti bioattivi richiede una selezione accurata della biomassa, privilegiando quella con una migliore stabilità nel tempo, ossia una migliore *shelf life*. (Gerardi *et al.*, 2022)

La *shelf life* di un prodotto è definita dalla norma ISO/TR 13097 come “il periodo di tempo raccomandato durante il quale un prodotto può essere conservato, nel corso del quale la qualità di una specifica proprietà rimane accettabile, in funzione delle condizioni previste (o specificate) di distribuzione, stoccaggio, esposizione e utilizzo.”

Comprendere le complesse interazioni tra fattori ambientali e fisiologia della biomassa vitivinicola è fondamentale per sviluppare strategie di conservazione efficaci, in grado di ridurre il degrado dei composti bioattivi e garantire la qualità del prodotto finale. (Yu *et al.*, 2024) Nel caso dell'uva e dei relativi scarti, la qualità della biomassa diminuisce rapidamente nel tempo a causa di processi come disidratazione, senescenza e degradazione dei composti fenolici. (Oren-Shamir *et al.*, 2009) Inoltre, con il progredire della maturazione e durante lo stoccaggio, aumenta l'attività di enzimi come le perossidasi, che contribuiscono alla degradazione delle antocianine. (Sha *et al.*, 2021) La perdita di umidità favorisce anche la crescita di muffe, con possibili effetti sulla sicurezza del prodotto a causa della produzione di micotossine. (Welke, 2019)

Segue nei sottoparagrafi 1.5.1-1.5.5 una rassegna di possibili indicatori da considerare per predire la *shelf life* di scarti vitivinicoli.

1.5.1 COMPOSTI FENOLICI E VARIAZIONI CROMATICHE

L'uva è costituita da centinaia di composti la cui concentrazione varia durante le diverse fasi di sviluppo dell'acino; i principali sono l'acqua (75–85%), seguita da zuccheri (glucosio, fruttosio e saccarosio), acidi organici (malico, tartarico e citrico), aminoacidi, proteine e composti fenolici (fenoli, antociani e flavonoli). (Jordão *et al.*, 2021)

Tra questi, i composti fenolici rappresentano i migliori indici per monitorare la qualità della biomassa vitivinicola, anche in considerazione del loro elevato potenziale applicativo. (Oancea, 2021) L'imbrunimento della biomassa e la variazione del colore sono infatti indicatori della degradazione dei composti bioattivi, che costituiscono al contempo la frazione più interessante da recuperare ed estrarre. (Muche *et al.*, 2018)

Sebbene le antocianine siano responsabili della colorazione rossa della vinaccia, la stabilità e l'intensità cromatica di quest'ultima dipendono dalle interazioni con polifenoli incolore. (Boulton *et al.*, 2001) Questi composti interagiscono con gli antociani e ne proteggono la struttura attraverso un fenomeno noto come copigmentazione, che avviene mediante un legame non covalente e porta alla formazione di un complesso di trasferimento di carica (π - π) tra molecole dotate di anelli aromatici con diversa densità elettronica. (Rodríguez-Pulido *et al.*, 2022) La modifica del colore dovuta a questo fenomeno produce deviazioni dalla legge di Lambert–Beer ed è stata ampiamente studiata non solo tramite misure di assorbanza a lunghezze d'onda specifiche, ma anche considerando l'intero spettro visibile e valutando il colore in modo olistico. (Gordillo *et al.*, 2012)

Ciononostante, le antocianine rappresentano la classe di fenoli più instabile, poiché suscettibili a numerose reazioni di degradazione, sia enzimatiche che non enzimatiche, come l'imbrunimento, la polimerizzazione e la condensazione con i tannini, durante le fasi di trasformazione e conservazione dei prodotti vitivinicoli. (Muche *et al.*, 2018) La loro stabilità cromatica dipende da molteplici fattori, quali pH, temperatura, tempo di conservazione, luce, ossigeno, concentrazione, acido ascorbico, copigmenti, enzimi, flavonoidi, proteine, ioni metallici e zuccheri. (Patras *et al.*, 2010, Muche *et al.*, 2018) Anche i flavonoli, sebbene più stabili rispetto agli antociani, risultano sensibili alle alte temperature, e il loro potenziale di estrazione tende a ridursi con l'aumento termico, influenzando così il contenuto fenolico complessivo e la qualità della materia prima. (Ramos *et al.*, 2024)

Pertanto, l'analisi dell'evoluzione dei composti fenolici, in particolare di antocianine e flavonoli, risulta fondamentale per comprendere i meccanismi di degradazione che

influenzano la stabilità chimica, il colore e la qualità complessiva della biomassa vitivinicola durante la conservazione. (Gerardi *et al.*, 2022)

1.5.2 LA TEMPERATURA

Un parametro importante che può influenzare la *shelf life* dei residui di produzione vitivinicola destinati all'estrazione di composti fenolici bioattivi è la temperatura. (Panza *et al.*, 2025) Le fluttuazioni termiche durante il trasporto e nei laboratori possono alterare significativamente l'attività metabolica della biomassa d'uva, influenzando processi fisiologici e biochimici come i tassi di respirazione, l'attività enzimatica e la degradazione dei componenti cellulari. (Duan *et al.*, 2020)

È stato infatti dimostrato che le diverse temperature di conservazione influenzano i tassi di respirazione dei frutti, incidendo così sulla loro *shelf life* e sulle caratteristiche sensoriali. (Jung *et al.*, 2018) Con l'aumento della respirazione i frutti, e analogamente i residui d'uva, vanno incontro a un'accelerata spesa energetica, che porta all'esaurimento delle riserve di amido o zuccheri, causando perdita di nutrienti e decadimento. (Yu *et al.*, 2022) La respirazione anaerobica, inoltre, intensifica il processo di invecchiamento, aumentando il consumo di nutrienti e l'attività dei processi metabolici necessari al mantenimento della vitalità cellulare. (Yu *et al.*, 2024)

Inoltre, l'impatto della temperatura si estende anche ai tassi di traspirazione: temperature elevate possono accelerare la perdita d'acqua, causando l'appassimento dei tessuti e influenzando negativamente la qualità della biomassa. (Cheng *et al.*, 2021) Pertanto, un accurato controllo della temperatura durante le fasi di trasporto e manipolazione è fondamentale per regolare i processi fisiologici, preservare la freschezza della biomassa e minimizzare il rischio di contaminazioni da parte di specie fungine. (Yu *et al.*, 2024)

1.5.3 UMIDITÀ RELATIVA

La biomassa di scarto vitivinicola, così come l'uva fresca appena raccolta, tende rapidamente a perdere acqua attraverso il processo di traspirazione, che rappresenta una delle principali cause di perdita di peso e deterioramento della struttura del materiale, compromettendo la qualità e l'integrità superficiale. (Jalali *et al.*, 2017) Questo fenomeno, particolarmente accentuato in ambienti caldi e secchi, dipende da caratteristiche specifiche dell'uva, quali il rapporto superficie/volume, la presenza di cere epicuticolari, la tenerezza della buccia e la

presenza di microlesioni generate durante i processi di vinificazione. (Sousa Gallagher *et al.*, 2011)

L'evaporazione dell'acqua avviene quando la pressione di vapore interna del tessuto vegetale supera quella dell'aria circostante, generando un gradiente che favorisce la perdita di umidità verso l'esterno. (Sousa Gallagher *et al.*, 2011) Tale processo è regolato dal deficit di pressione di vapore (VPD), che rappresenta la differenza tra la pressione di vapore alla saturazione e quella effettiva dell'aria. (Fanourakis *et al.*, 2013) Sebbene il VPD offra una misura più accurata della domanda evaporativa dell'aria, nella pratica orticola si utilizza più frequentemente l'umidità relativa. (Fanourakis *et al.*, 2013)

Condizioni ambientali caratterizzate da bassa umidità relativa e presenza di vento, come nell'essiccazione naturale delle bacche, favoriscono fortemente la disidratazione. (Sanmartin *et al.*, 2021)

L'eccessiva evaporazione d'acqua stimola inoltre l'attività di idrolasi come clorofilasi e pectinasi, accelerando la degradazione di composti legati alla qualità, come clorofilla e pectina, con conseguenti alterazioni nella consistenza, nel colore e nel sapore, e un accumulo di solidi solubili. (Moggia *et al.*, 2017) Infine, la disidratazione può anche innescare la sintesi e accelerare il rilascio di etilene, contribuendo ulteriormente al deterioramento della biomassa, in modo analogo a quanto osservato nei frutti freschi. (Yu *et al.*, 2024)

Una corretta gestione della temperatura e dell'umidità relativa è essenziale per ridurre la disidratazione dal momento che un eccesso di umidità può favorire lo sviluppo di muffe, mentre un'umidità insufficiente accelera la disidratazione. (Sousa Gallagher *et al.*, 2011)

1.5.4 VOCs

La concentrazione e la composizione dei composti organici volatili (VOCs) rappresentano un indicatore utile dello stadio di maturazione dell'uva, poiché la loro distribuzione percentuale varia significativamente tra le fasi fenologiche di allegagione, invaiatura, maturazione e post-maturazione. (Liu *et al.*, 2022) Queste stesse caratteristiche, tuttavia, possono essere sfruttate anche per stimare la *shelflife* delle bucce d'uva residue post-vinificazione, poiché il contenuto e il profilo dei VOCs riflettono lo stato fisiologico dell'uva al momento della lavorazione, che dipende in larga misura dall'attività e specificità enzimatica, come quella di alcool deidrogenasi, alcool acetil transferasi e enal isomerasi, coinvolte nella biosintesi e trasformazione dei composti aromatici. (Kalua *et al.*, 2009)

Sebbene sia difficile generalizzare a causa delle differenze varietali tra uve rosse, comprese quelle neutrali e semi-aromatiche, è stato dimostrato, analizzando separatamente la polpa e

la buccia, che fattori come la temperatura, il grado di disidratazione e la conseguente perdita di peso influenzano significativamente il profilo dei VOCs. (Piombino *et al.*, 2025)

Inoltre, alcuni composti volatili sono stati rilevati esclusivamente nelle fasi avanzate di maturazione in grappoli d'uva rossa schiacciati e nel relativo mosto fermentato, tra i quali bourbonene, (*E*)-geranil-acetone, etil-esanoato, estere (*Z*)-acido butanoico, 3-esenile acetato ed estere etilico dell'acido dodecanoico, suggerendo un legame diretto tra avanzamento della maturazione e arricchimento in specifici VOCs. (Ascirizzi *et al.*, 2022)

In questo contesto, l'analisi del profilo dei VOCs, con particolare attenzione alla frazione glicosilata nelle vinacce, potrebbe rappresentare un utile indicatore dello stato biochimico e del grado di degradazione della biomassa, nonché un potenziale predittore della sua *shelf life* e della conservazione di composti bioattivi come gli antociani. (Piombino *et al.*, 2025)

1.5.5 CONTAMINAZIONI MICROBIOLOGICHE E MICOTOSSINE

Durante la conservazione della biomassa vitivinicola di scarto, destinata all'estrazione di composti fenolici o ad altri impieghi, è essenziale considerare il rischio di contaminazione microbiologica da parte di specie fungine e batteriche, oltre alla possibile presenza di micotossine. (Aning-Dei *et al.*, 2025)

Temperature elevate o un eccesso di umidità disponibile possono favorire lo sviluppo di muffe e lieviti, accelerando fenomeni di marcescenza, fermentazioni indesiderate e degradazione generale della biomassa. (Sousa Gallagher *et al.*, 2011; Izquierdo-Bueno *et al.*, 2024) Numerose evidenze dimostrano che i microrganismi sono in grado di sopravvivere anche in condizioni ambientali sfavorevoli, come basse temperature, purché sia presente acqua in forma liquida. (Alegbeleye *et al.*, 2022) Per questo motivo, molte strategie di conservazione degli alimenti si basano sul controllo dell'attività dell'acqua (*aw*) del prodotto o dell'ambiente circostante. (King *et al.*, 2017) In particolare, l'umidità residua superficiale rappresenta un fattore critico nella proliferazione microbica. (Barth *et al.*, 2009)

La biomassa vitivinicola è spesso soggetta alla contaminazione da parte di muffe micotossigene, in particolare *Aspergillus* e *Penicillium*, in grado di produrre ocratossina A (OTA), una micotossina frequentemente rilevata nelle vinacce. (Solfrizzo *et al.*, 2008; Yu *et al.*, 2018; Yu *et al.*, 2020) L'OTA rappresenta un grave rischio per la salute umana e animale: diversi studi ne documentano la nefrotossicità, epatotossicità, immunotossicità, neurotossicità, teratogenicità e potenziale cancerogenicità. (Yu *et al.*, 2023)

In ambito zootecnico, l'impiego di mangimi contaminati da OTA ha compromesso le *performance* degli animali da allevamento e favorito l'accumulo residuo della tossina nelle

carni, con conseguenti rischi per la sicurezza alimentare dei consumatori. (Yu *et al.*, 2023)
Pertanto, la presenza di OTA o di altre micotossine limita fortemente la possibilità di valorizzare le vinacce, sia in applicazioni commestibili, sia come fonte di composti bioattivi destinati all'uso umano o animale. (Aning-Dei *et al.*, 2025)

1.6 EVOLUZIONE DEI METODI PREDITTIVI DELLA *SHELF LIFE* DI BIOMASSA DEPERIBILE

A livello operativo, la determinazione della *shelf life* di un prodotto alimentare si basa generalmente sul monitoraggio di specifici parametri di qualità che variano nel tempo. (Chaudry *et al.*, 2018)

La *shelf life* viene comunemente stimata attraverso due diverse tipologie di test di stabilità: i test in tempo reale, in cui il prodotto viene conservato alle condizioni di stoccaggio raccomandate e monitorato fino al superamento delle specifiche di qualità, e i test di stabilità accelerata, in cui il prodotto è sottoposto a condizioni di stress elevate (come temperatura, umidità e pH), durante le quali vengono valutati i cambiamenti chimici, biochimici e microbiologici. (Haouet *et al.*, 2019)

L'approccio *Accelerated Shelf Life Testing* (ASLT), in particolare, risponde all'esigenza di valutare in tempi brevi la stabilità e la qualità di un prodotto destinato alla commercializzazione. (Cui *et al.*, 2023) In questo caso, i campioni vengono conservati in condizioni accelerate (temperature elevate, maggiore umidità, esposizione all'ossigeno, ecc.) per simulare e velocizzare i processi di degradazione, e i dati ottenuti vengono successivamente utilizzati per costruire modelli cinetici che consentono di stimare la stabilità e la conservabilità del prodotto in condizioni di conservazione reali. (Hough *et al.*, 2006)

Dal momento che la tecnica ASLT è di tipo monovariato e quindi inadeguato a descrivere fenomeni complessi come la degradazione di biomasse vegetali complesse, si è scelto di integrare strumenti statistici multivariati, come quelli offerti dalla chemiometria, portando allo sviluppo del *Multivariate Accelerated Shelf Life Testing* (MASLT), un metodo che consente di monitorare simultaneamente più attributi qualitativi, offrendo una stima più completa e affidabile della *shelf life* di diversi prodotti alimentari. (Chaudhry *et al.*, 2018)

Nel contesto della chemiometria, il MASLT è spesso associato all'Analisi delle Componenti Principali (PCA), che consente di ridurre la dimensionalità dei dati, identificare *pattern* significativi tra le variabili e attenuare il rumore sperimentale, oppure alla regressione PLS (*Partial Least Squares Regression*, PLSR), che permette di modellare direttamente la relazione

tra la matrice dei dati e le proprietà chimiche, ricercando componenti latenti capaci di massimizzare la covarianza tra le variabili predittive e quelle di risposta. (Lima *et al.*, 2023)

Tuttavia, i modelli matematici tradizionali si basano spesso su ipotesi semplificate che non riescono a cogliere appieno la complessità dei meccanismi di deterioramento, portando a potenziali imprecisioni nelle applicazioni del mondo reale. (Rashvand *et al.*, 2025) In questo contesto di sfide, l'intelligenza artificiale (IA) è emersa come una potenziale soluzione. (Shi *et al.*, 2023)

I metodi basati sull'IA, attraverso il *machine learning* (ML) e il *deep learning* (DL), sono in grado di analizzare grandi volumi di dati e individuare schemi difficilmente rilevabili dai modelli tradizionali, offrendo inoltre la possibilità di includere molteplici tipologie di variabili, come la composizione biochimica, le condizioni di conservazione e l'attività microbica. (Rashvand *et al.*, 2025)

Per poter addestrare un modello di intelligenza artificiale alla previsione della *shelf life* di biomassa vitivinicola, è fondamentale acquisire una conoscenza approfondita degli indicatori fisiologici e biochimici che caratterizzano l'uva durante lo sviluppo e la post-raccolta, selezionando variabili significative e diversificate, essenziali per costruire modelli predittivi affidabili. (Xu *et al.*, 2024)

1.7 MODELLI DI INTELLIGENZA ARTIFICIALE PER LA *SHELF LIFE* DI BIOMASSA VITIVINICOLA

Già nel 2012 il gruppo di Rodríguez-Pulido *et al.* (Rodríguez-Pulido *et al.*, 2012a) ha presentato il primo studio volto a valutare il potenziale della visione artificiale nella determinazione della maturità fenolica dei semi d'uva, attraverso un'analisi di regressione lineare univariata, seguita da un'analisi multivariata. I risultati hanno evidenziato correlazioni significative fra la composizione fenolica e l'aspetto visivo (colore e morfologia) dei semi, acquisendo i dati con una semplice telecamera digitale. Nel medesimo anno, lo stesso gruppo di ricerca ha esteso l'analisi anche agli acini, determinando la dimensione e lo stadio di sviluppo degli acini attraverso tecniche di *image analysis*. (Rodríguez-Pulido *et al.*, 2012b) Nel 2014 Avila *et al.* (Avila F. *et al.*, 2014) hanno presentato una tecnica ibrida, combinando approcci supervisionati e non supervisionati, per classificare i semi d'uva come maturi o immaturi. Il loro metodo si basa su immagini di semi utilizzando modelli di colore invarianti e una rete neurale perceptrone multistrato (*Multilayer Perceptron*, MLP).

Nello stesso anno, il gruppo di Zuñiga *et al.* (Zuñiga *et al.*, 2014) ha ampliato la classificazione in tre classi (immaturi, maturi, oltre-maturi), utilizzando tre MLPs in parallelo.

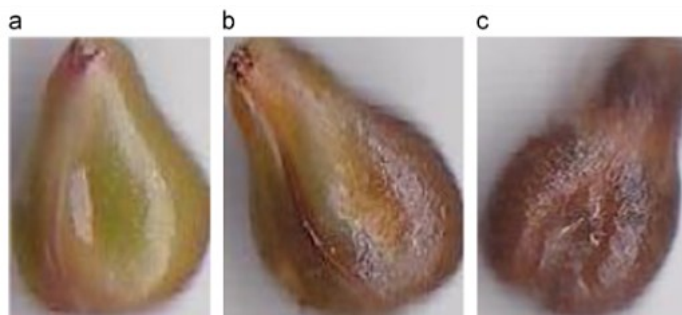


Figura 1.4. Classificazione dei semi in base al grado di maturità: immaturo, maturo, oltre-maturo. (da Zuñiga *et al.* 2014)

Nel 2022, Navarro P. J. *et al.* (Navarro *et al.*, 2022) hanno testato modelli *MLP* e *3D Convolutional Neural Networks (3D-CNN)* su un dataset multispettrale di acini d'uva e contenente misure di peso, antociani e indice Brix. Più recentemente, nel 2024 Xu M. *et al.* (Xu *et al.*, 2024) hanno impiegato l'*imaging* iperspettrale Vis/NIR per stimare in modo non distruttivo la *shelf life* dell'uva da tavolo di Kyoho utilizzando un modello ibrido di *deep learning* SDAE-SVM, confrontato con altre architetture quali PCA-SVM, LPP-SVM, *one-dimensional convolutional neural network* (1D CNN), *long short-term memory* (LSTM), 1D CNN-SVM, LSTM-SVM.

Uno degli esempi più recenti di integrazione tra *imaging* multispettrale e *deep learning* avanzato per la valutazione non distruttiva della qualità dell'uva è stato presentato nel 2025 da (Chen *et al.*, 2025), che hanno proposto un modello basato su rete ResNet-50, denominato *Improved-Res*, per stimare in modo non distruttivo il contenuto zuccherino ($^{\circ}$ Brix) dell'uva *Sunshine Rose*, ottenendo prestazioni elevate (MSE = 0.49, MAE = 0.55 $^{\circ}$ Brix, $R^2 = 0.92$) e dimostrando l'efficacia dei moduli di attenzione e delle reti residuali nella predizione dei parametri qualitativi dei frutti.

Nel 2023, Tsakiridis N. L. *et al.* (Tsakiridis *et al.*, 2023) hanno esplorato architetture *autoencoders* (*Denosing Autoencoder*, DAE e *Convolutional Autoencoder*, DCAE) in combinazione con algoritmi regressivi di *machine learning* (*Random Forest*, *Cubist*, *XGBoosting* e *Support Vector Regression*, SVR), al fine di predire il grado di maturazione di diverse varietà di uva. Già nel 2020, alcuni membri del gruppo di ricerca Tsakiridis N. L. *et al.* (Samarinas *et al.*, 2020) avevano già testato *Random Forest* e *Partial Least Squares (PLS)* su dati spettroscopici VNIR–SWIR per predire la maturazione dell'uva *Syrah*.

Nel novembre 2023, il gruppo di Kalopesa *et al.* (Kalopesa *et al.*, 2023), ha testato diversi modelli di *machine learning*, tra cui *Random Forest*, *Partial Least Squares*, *Support Vector Regression*, *Convolutional Neural Network* e un modello di reti neurali convoluzionali *multi-input–multi-output* da loro sviluppato, per stimare la maturazione dell’uva da vino utilizzando la spettroscopia infrarossa.

Nello stesso anno, il gruppo di ricerca di Blekos A. *et al.* (Blekos *et al.*, 2023) ha pubblicato il *CERTH Grape Dataset*, utilizzato per la stima della maturità del grappolo mediante algoritmi di segmentazione istanza-based come *Mask2Former Swin (S)*, che ha mostrato le migliori *performance* nel riconoscimento multi-classe.

Infine, nel 2019 il gruppo di Li Y. *et al* (Li *et al.*, 2019) ha sviluppato una nuova rete RBF come modello di previsione della *shelf life* residua (*Residual Shelf Life*) per l’uva da tavolo post-raccolta, considerando l’uva bianca italiana e l’uva rossa Red Globe.

In sintesi, dispositivi di rilevamento a contatto e non a contatto, come telecamere ad alta definizione e spettrometri VIS–NIR, combinati con modelli di *machine learning* e reti neurali profonde, costituiscono oggi lo stato dell’arte per la valutazione della qualità e della *shelf life* dell’uva. (Chen *et al.*, 2025)

Tabella II. Riepilogo degli articoli analizzati nella ricerca bibliografica

Autore (anno)	Obiettivo dello studio	Tecniche utilizzate
Rodríguez-Pulido et al., 2012a	Determinare la maturità fenolica dei semi d'uva mediante la correlazione fra composizione fenolica e aspetto visivo	Regressione lineare univariata + analisi multivariata; image analysis
Rodríguez-Pulido et al., 2012b	Stima della maturazione di acini e semi	<i>Image analysis</i> RGB
Avila et al., 2014	Classificazione binaria dei semi d'uva (maturi/immaturo) da immagini RGB	Segmentazione ibrida (supervisionata e non), MLP
Zuñiga et al., 2014	Classificazione dei semi in 3 classi di maturità (maturi/immaturo/oltre-maturi)	Tre MLP, modello colore invariante
Navarro et al., 2022	Predizione del grado di maturazione degli acini	MLP e 3D-CNN, dataset multispettrale
Xu et al., 2024	Stima non distruttiva della shelf-life dell'uva Kyoho	Vis/NIR imaging + deep learning (SDAE-SVM, LSTM, 1D CNN)
Chen et al., 2025	Stima non distruttiva della shelf-life dell'uva <i>Sunshine Rose</i>	VIS/NIR imaging multispettrale + rete residuale- ResNet50
Tsakiridis et al., 2023	Predizione maturità uva da immagini iperspettrali	DAE, DCAE + RF, Cubist, XGBoost, SVR
Samarinas et al., 2020	Stima maturazione uva Syrah con spettroscopia VNIR-SWIR	RF, PLS
Kalopesa et al., 2023	Predizione del livello di maturazione da parametri chimici, utilizzando spettroscopia IR	CNN, MIMO-CNN, RF, PLS, SVR
Blekos et al., 2023	Classificazione del livello di maturità dei grappoli da immagini RGB	Mask R-CNN, Mask2Former, Cascade Mask R-CNN, HTC-R101
Li et al., 2019	Predizione della shelf-life dell'uva post-raccolta	RBF Neural Network ottimizzata

1.8 MODELLI DI INTELLIGENZA ARTIFICIALE UTILIZZATI IN QUESTO STUDIO

1.8.1 REGRESSIONE LINEARE MULTIPLA (MLR)

La regressione lineare multipla (**Figura 1.5**) rappresenta un'estensione naturale della regressione lineare semplice: invece di relazionare una variabile dipendente (y) a una singola variabile indipendente (x), si cerca di spiegare (y) come somma ponderata degli effetti di più variabili indipendenti (x_1, x_2, x_3, \dots). In forma matematica:

$$y = k + ax_1 + bx_2 + cx_3 + \dots + \varepsilon,$$

dove k è l'intercetta sull'asse delle y , a, b, c sono i coefficienti che quantificano l'effetto delle variabili indipendenti, x_1, x_2, x_3 etc, sono le diverse *features* che si usano per spiegare o prevedere la variabile y , e ε rappresenta il rumore o l'errore casuale. (Smalheiser, 2017)

Rispetto alla regressione semplice, il modello viene interpretato come una retta di miglior adattamento nello spazio a $(n + 1)$ dimensioni, dove (n) è il numero delle variabili indipendenti. (Smalheiser, 2017)

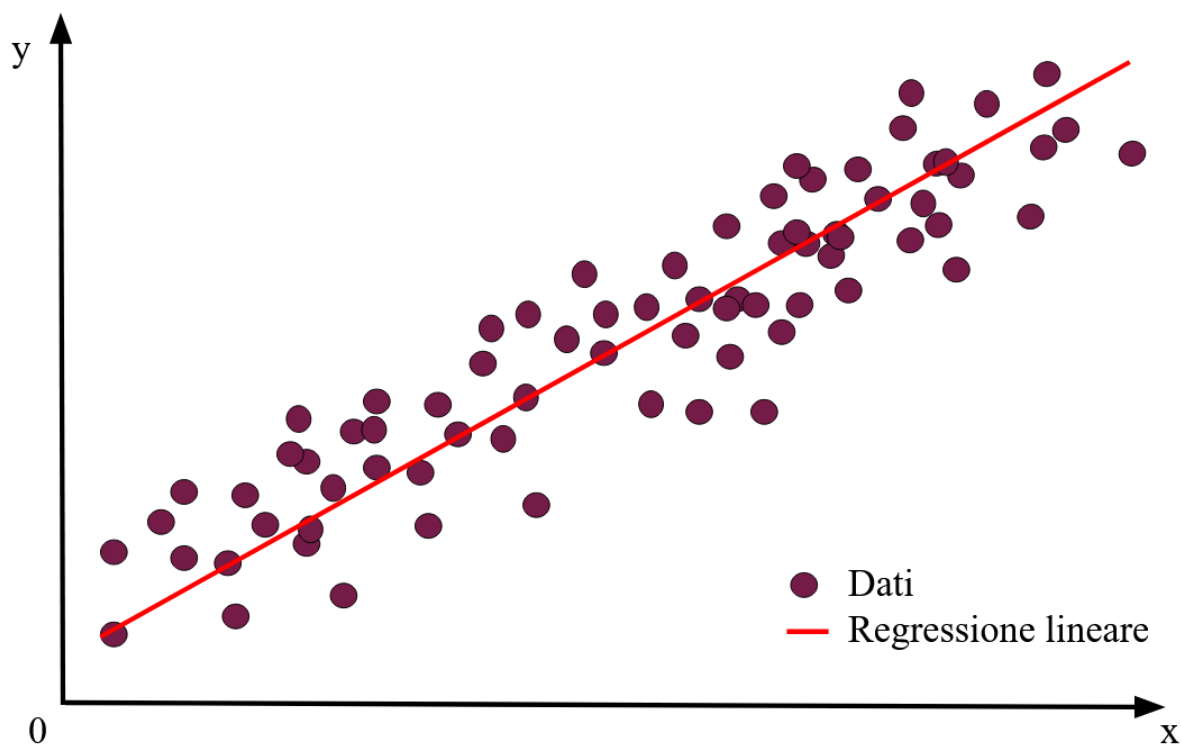


Figura 1.5. Rappresentazione della regressione lineare multipla.

1.8.2 RANDOM FOREST REGRESSOR (RFR)

Il modello *Random Forest Regressor* è un approccio *ensemble* costituito da diversi alberi decisionali indipendenti e non correlati, costruiti in parallelo, ciascuno dei quali viene addestrato su un sottoinsieme casuale del dataset e delle variabili o *features* (**Figura 1.6**). (Singh *et al.*, 2021)

Ogni albero viene infatti addestrato su un campione bootstrap (vedi Paragrafo 1.9), ottenuto estraendo con reinserimento n osservazioni dal dataset originale, con possibili duplicati, e durante la costruzione degli *split* dei nodi degli alberi decisionali, selezionando casualmente un sottoinsieme di k *features*, con $k < p$ dove p rappresenta il numero totale di variabili disponibili. (Ha, 2023)

Gli alberi così generati, tra loro debolmente correlati, vengono successivamente combinati tramite un meccanismo di aggregazione: nella classificazione si utilizza la maggioranza delle predizioni dei singoli alberi, mentre nella regressione la stima finale è ottenuta come media delle loro predizioni. (Salman *et al.*, 2024)

Grazie all'introduzione di casualità nella loro costruzione e all'impiego della tecnica di *bagging* (che combina *bootstrapping* e aggregazione), il modello consente di ottenere una buona generalizzazione, risultando particolarmente efficace nella gestione di relazioni non lineari tra le variabili, riducendo la varianza e migliorando la capacità predittiva rispetto a singoli alberi decisionali. (El Mbrabet *et al.*, 2022)

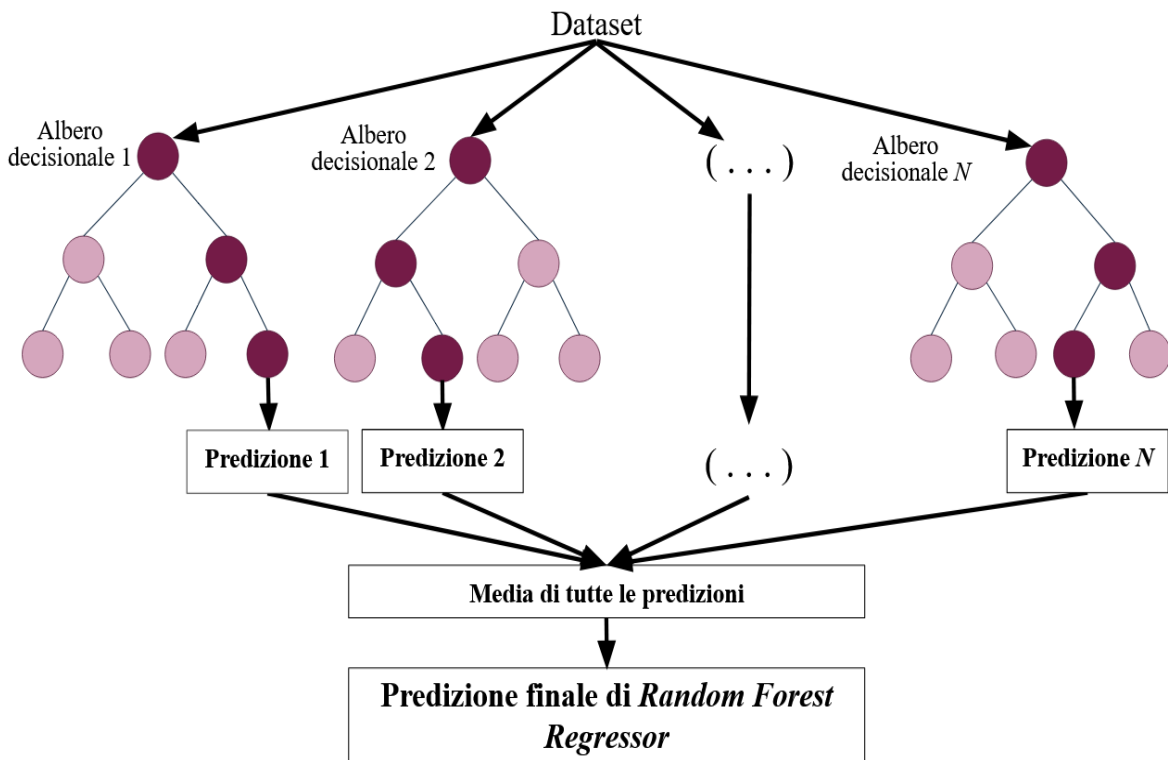


Figura 1.6. Rappresentazione dell’algoritmo *Random Forest Regressor*. Il modello costruisce molteplici alberi decisionali addestrati su differenti sottoinsiemi del dataset e delle variabili; ogni albero produce una propria predizione e il risultato finale deriva dalla combinazione delle predizioni dei singoli alberi, tramite media nel caso della regressione.

1.8.3 GRADIENT BOOSTING REGRESSOR (GBR)

Il modello *Gradient Boosting Regressor* è un metodo di *ensemble learning* che combina sequenzialmente più alberi decisionali (*weak learners*), ognuno dei quali corregge gli errori residui commessi dai precedenti attraverso un processo iterativo di ottimizzazione basato sulla discesa del gradiente (**Figura 1.7**). (Singh *et al.*, 2021)

A ogni iterazione, il modello apprende una nuova funzione che approssima i residui dell’iterazione precedente, ovvero la componente dell’errore non ancora spiegata, e questo principio è alla base del *boosting* sia in compiti di regressione sia di classificazione: ogni nuovo modello debole viene addestrato per migliorare progressivamente le prestazioni dell’insieme, correggendo gli errori residui non ancora catturati. (Friedman, 2001)

Nei modelli di *Gradient Boosting*, gli alberi decisionali sono combinati in modo additivo per formare un modello finale più accurato, in cui l’insieme dei *weak learners* dà origine a un modello “forte” capace di catturare *pattern* complessi nei dati, impiegando, a seconda del *task*,

funzioni di perdita specifiche come l'errore quadratico medio per la regressione o la perdita logaritmica per la classificazione. (Deng *et al.*, 2021)

Questo approccio additivo consente di ridurre progressivamente il *bias* e di catturare relazioni non lineari tra le variabili, migliorando la capacità predittiva e la generalizzazione del modello. (Singh *et al.*, 2021)

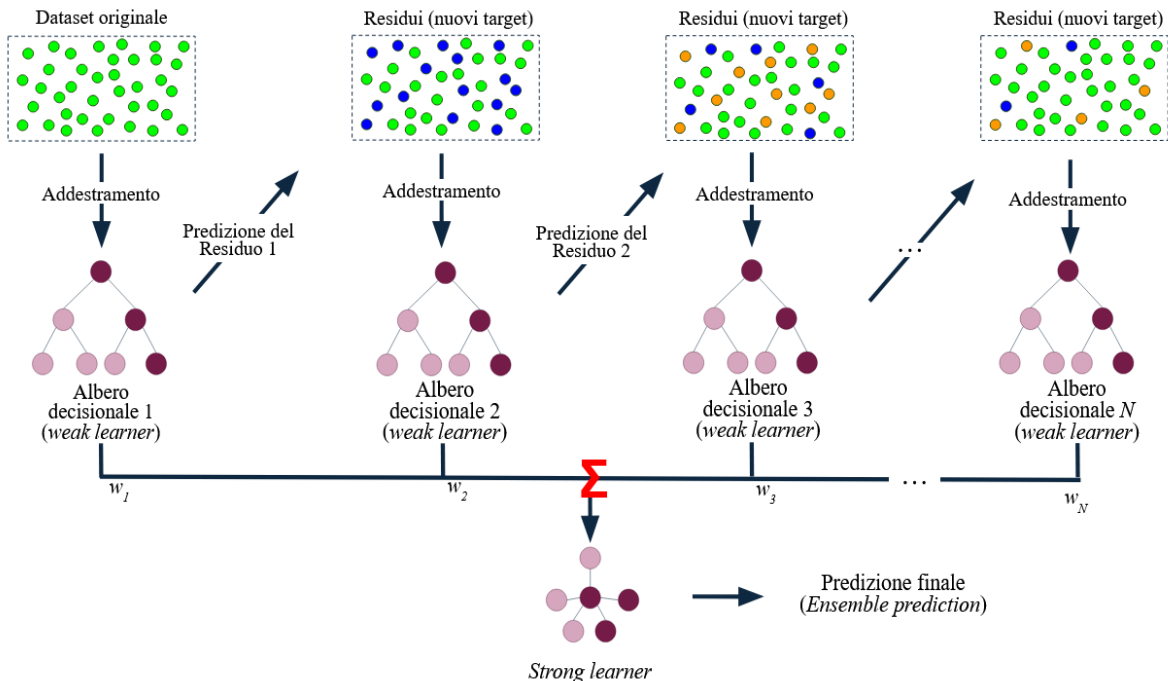


Figura 1.7. Schema del funzionamento del *Gradient Boosting Regressor*. Il modello inizia dall'addestramento sul dataset originale (pallini verdi). A ogni iterazione vengono calcolati i residui (pallini blu, arancioni), e un albero decisionale viene addestrato per approssimarli. I modelli vengono poi combinati in modo additivo (pesi $w_1, w_2, w_3, \dots, w_N$) e dopo N iterazioni, l'insieme dei *weak learners* forma uno *strong learner* che produce la predizione finale.

1.8.4 SUPPORT VECTOR REGRESSION (SVR)

La *Support Vector Regression* (SVR), derivata dall'approccio delle *Support Vector Machines* (SVM), è una tecnica di *machine learning* supervisionato per la predizione di variabili continue, che individua una funzione $f(x)$ in grado di approssimare i valori osservati entro una soglia di errore prefissata ϵ , minimizzando al contempo la complessità del modello, ovvero massimizzandone la "piattezza" (Figura 1.8). (Rohani *et al.*, 2018)

Questo approccio consente alla SVR di catturare in modo flessibile le relazioni non lineari tipiche dei processi di degradazione alimentare, nei quali l'interazione tra fattori quali temperatura, umidità, pH e luce determina dinamiche complesse di alterazione microbiologica

e ossidativa. (Namazi *et al.*, 2025) La capacità di modellare tali non linearità è ottenuta tramite l'impiego di funzioni kernel, ovvero funzioni matematiche che mappano implicitamente i dati in spazi di dimensione superiore, evitando così i limiti dei modelli polinomiali classici spesso soggetti a *underfitting* o *overfitting*. (Alizadeh Behbahani *et al.*, 2024)

Seguendo quanto descritto in letteratura, i dati possono essere concettualmente rappresentati come punti nello spazio delle caratteristiche, mentre la funzione stimata dalla SVR forma una curva o superficie continua che approssima la variabile target, evidenziando il comportamento del modello in relazione ai dati osservati. (Awad *et al.*, 2015)

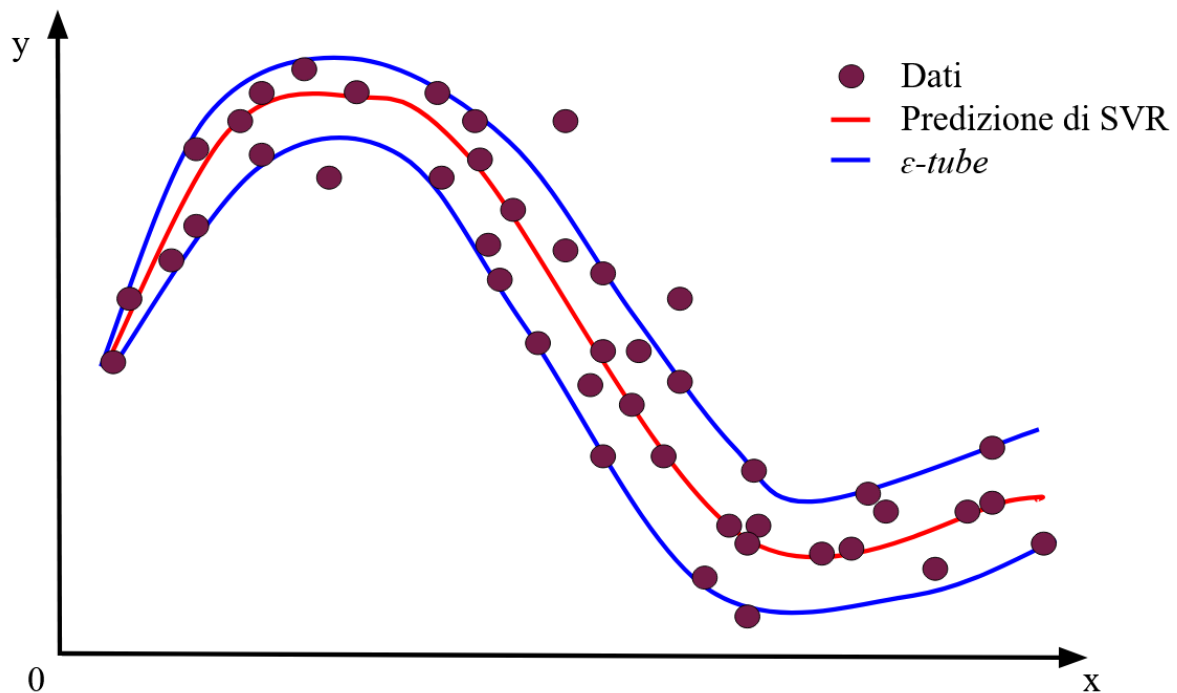


Figura 1.8. Schema del funzionamento della *Support Vector Regression (SVR)*. La funzione $f(x)$ (linea rossa) approssima i valori osservati (sfere viola) entro una soglia di errore prefissata ϵ (linee blu), mantenendo al minimo la complessità del modello.

1.8.5 RETI NEURALI ARTIFICIALI (*MULTILAYER PERCEPTRON, MLP*)

Il modello *Multilayer Perceptron (MLP)* è una rete neurale artificiale comunemente utilizzata per compiti di apprendimento supervisionato, inclusa la regressione, composta da uno strato di input, uno o più strati nascosti che applicano trasformazioni non lineari alle feature, e uno strato di output che produce la predizione finale (**Figura 1.9**). (Jain *et al.*, 1996)

Ogni strato è composto da neuroni interconnessi che ricevono gli input dallo strato precedente e li elaborano applicando una trasformazione lineare tramite pesi e termini di bias, seguita da

una funzione di attivazione (lineare o non lineare a seconda dello strato), trasmettendo così l'output allo strato successivo. (Abdelbasset *et al.*, 2022)

Per ottenere prestazioni ottimali è necessario regolare iperparametri come la funzione di attivazione, l'algoritmo di ottimizzazione (*solver*) e la dimensione degli strati nascosti, e l'output del modello può essere espresso come una composizione di funzioni attivate su più strati, riflettendo la struttura multilivello della rete. (Sumayly *et al.*, 2023)

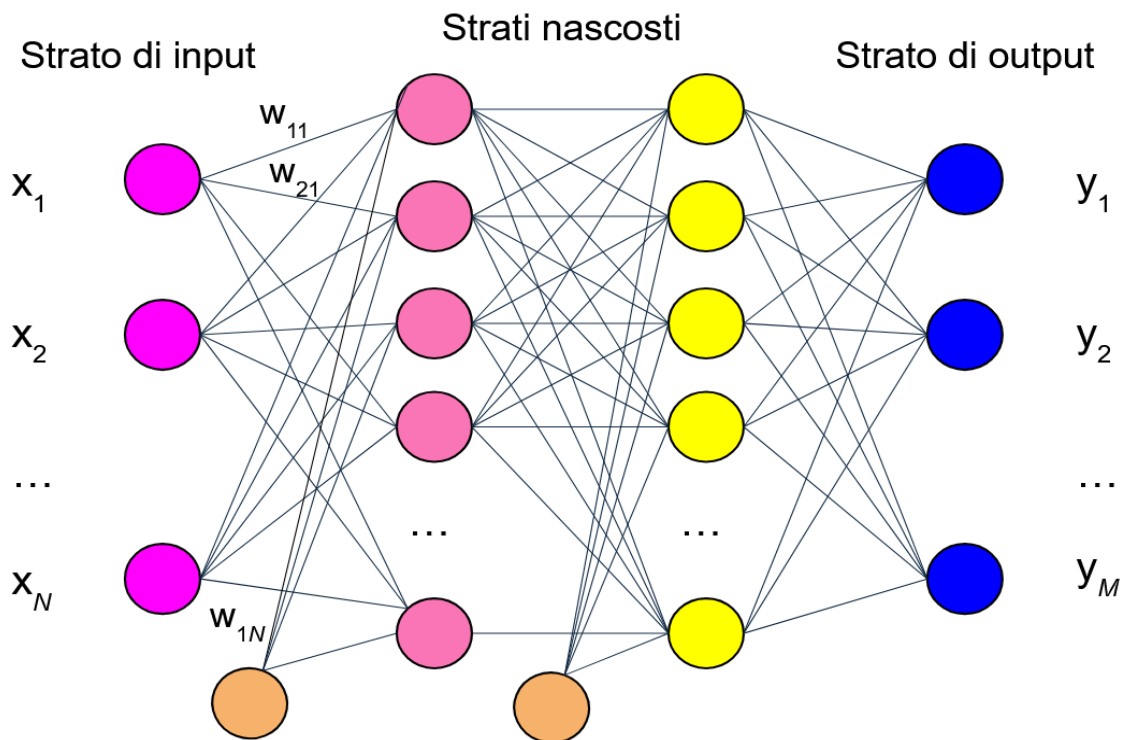


Figura 1.9. Schema del *Multilayer Perceptron (MLP)* con due strati nascosti. Lo strato di input (cerchi fucsia) trasmette i segnali agli strati nascosti (cerchi rosa e gialli) e allo strato di output (cerchi blu). Le connessioni hanno pesi w e ogni neurone include un *bias* (cerchi arancioni, per comodità visiva è stato rappresentato un neurone per strato) che regola la soglia della funzione di attivazione.

1.8.6 RETI NEURALI *DEEP LEARNING (RESIDUAL NETWORK, RESNET)*

La *Residual Network (ResNet)* è un'architettura di rete neurale convoluzionale 1D che utilizza blocchi residui per facilitare l'addestramento di reti profonde, mitigando il problema del gradiente evanescente e migliorando la capacità di apprendere rappresentazioni complesse dei dati (**Figura 1.10**). (Li *et al.*, 2025)

Strutturalmente, una ResNet è composta da una fase iniziale di estrazione delle caratteristiche (*embedding block*), seguita da una sequenza di blocchi residui che trasformano progressivamente lo spazio delle rappresentazioni (*mapping block*) realizzando una mappatura del tipo $y=F(x)+x$; infine, uno strato di predizione *fully connected* mappa la rappresentazione ottenuta nello spazio di output per la regressione o la classificazione (*prediction block*). (Rousseau *et al.*, 2020)

Il principio chiave consiste nell'introdurre connessioni di scorciatoia (*skip-connections*), che consentono al gradiente di propagarsi direttamente agli strati più profondi, stabilizzando il processo di addestramento, mentre ogni blocco residuo è costituito dalla concatenazione di layer di normalizzazione batch (BatchNorm), funzioni di attivazione ReLU e *layer* convoluzionali. (López García-Arias *et al.*, 2023)

I dati multivariati e sequenziali, con complesse interazioni non lineari, rendono difficile l'estrazione di informazioni pertinenti, ma CNN 1D e ResNet 1D consentono di apprendere rappresentazioni latenti efficaci dai dati grezzi, migliorando la capacità predittiva dei modelli supervisionati. (Tchatchoua *et al.*, 2023)

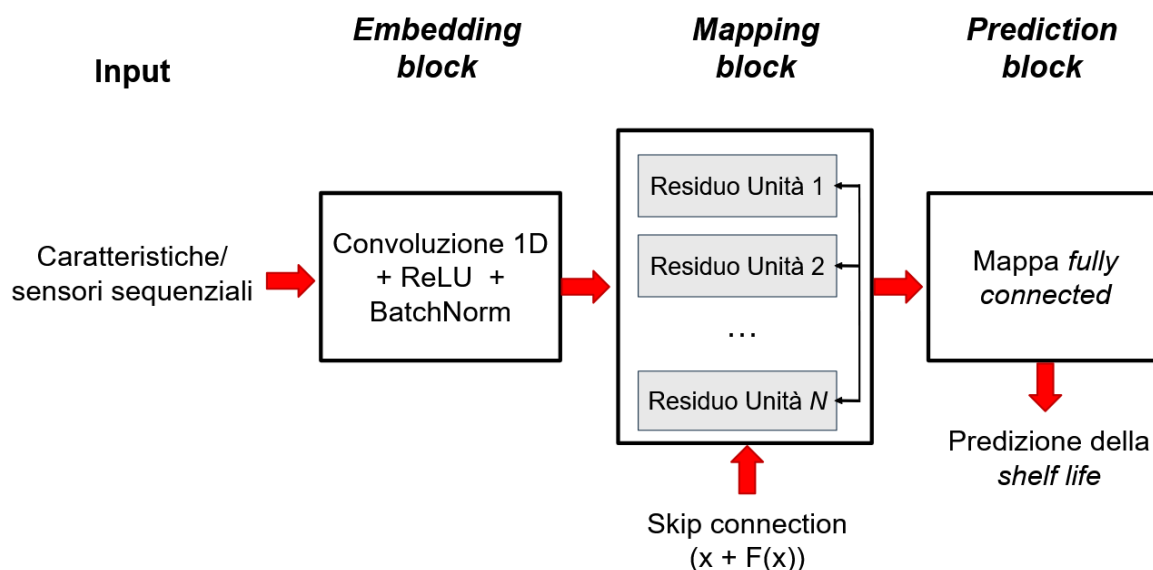


Figura 1.10. Schema della Residual Network (ResNet) 1D. I dati di input, come segnali di sensori o *features*, vengono elaborati dal blocco di estrazione (*embedding block*) con layer convoluzionali, ReLU e BatchNorm. Successivamente passano ai blocchi residui (*mapping block*), in cui le *skip-connections* sommano l'input x alla funzione appresa $F(x)$, facilitando la propagazione del gradiente. Infine, il blocco di predizione (*prediction block*) produce l'output finale per regressione o classificazione.

1.9 BOOTSTRAPPING

Nell'analisi di modelli di *machine learning*, le prestazioni del modello vengono spesso valutate mediante test basati su un singolo sottoinsieme di dati, che possono risultare poco robusti e non pienamente rappresentativi delle caratteristiche rilevanti del fenomeno osservato; inoltre, riservare parte dei dati alla verifica riduce le informazioni disponibili per l'addestramento, rendendo necessario interrogarsi sulla reale affidabilità delle misure di performance ottenute e sulle eventuali differenze tra modelli o dataset. (Koźniewski *et al.*, 2016)

Tradizionalmente, la valutazione dell'accuratezza di uno stimatore si basa sul calcolo dell'errore standard, ma per molti stimatori diversi dalla media non esiste una formula analitica che lo permetta. (Efron *et al.*, 1994) Per superare questi limiti, si ricorre al *resampling*, cioè estrarre ripetutamente campioni randomizzati da un dataset e riaddestrare il modello su ciascun campione, in modo che il tasso di errore medio su molteplici esecuzioni fornisca una stima più affidabile rispetto a una singola stima puntuale; tra i metodi di *resampling* più comuni si annoverano il Monte Carlo, il bootstrap, la *k-fold cross-validation* e la *repeated k-fold Cross-Validation*. (Nakatsu, 2023)

In particolare, il bootstrap è un metodo di ricampionamento per l'inferenza statistica, comunemente utilizzato per stimare intervalli di confidenza, ma anche per valutare il *bias* e la varianza di uno stimatore o calibrare test di ipotesi, poiché consente di stimare l'accuratezza delle stime statistiche anche con dataset limitati e quando la distribuzione dei dati è ignota. (Dixon, 2006)

Come illustrato in **Figura 1.11**, il *bootstrapping* consiste nel generare N campioni bootstrap dal dataset originale tramite campionamento con reinserimento, calcolare la statistica di interesse $s(X)$ su ciascun campione per ottenere le repliche $s(X^{*1}), s(X^{*2}), \dots, s(X^{*N})$ e stimare l'errore standard della statistica come deviazione standard delle repliche. (Efron, 1994) In questo modo, utilizzando la distribuzione empirica del campione come sostituta di quella della popolazione, il bootstrapping semplifica l'analisi di statistiche o test complessi e permette di valutare la variabilità e la precisione delle stime anche con dataset limitati. (Li *et al.*, 2013)

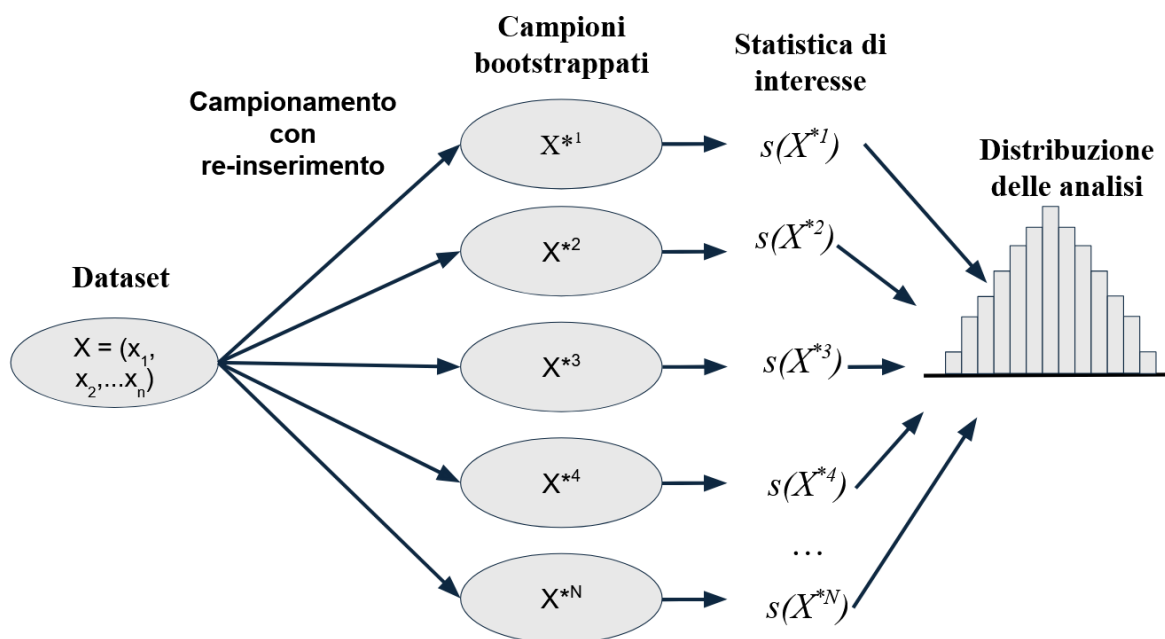


Figura 1.11. Schema del *bootstrapping*. Dal dataset originale $X=(x_1,x_2,\dots,x_n)$ vengono generati N campioni *bootstrap* ($X^{*1},X^{*2},\dots,X^{*N}$) tramite campionamento con reinserimento. La statistica di interesse $s(X)$ viene calcolata su ciascun campione, ottenendo $s(X^{*1}),s(X^{*2}),\dots,s(X^{*N})$. Infine, si costruisce la distribuzione delle statistiche calcolate sui diversi campioni.

1.10 SENSORI E IOT

Per predire con maggiore precisione la *shelf life* degli alimenti e rilevare in modo accurato le dinamiche di deterioramento della biomassa organica, i modelli di intelligenza artificiale possono analizzare dati provenienti da sensori ambientali di diversa tipologia, immagini, segnali spettrali e parametri di conservazione. (Shehzad *et al.*, 2025)

Questi sensori, integrati con tecnologie IoT (*Internet of Things*), inviano costantemente dati in tempo reale a piattaforme IA basate su cloud, dove vengono elaborati per fornire previsioni e rilevare automaticamente segnali di degradazione (Shehzad *et al.*, 2025)

Più in dettaglio, l'*Internet of Things* è costituito da dispositivi e macchine dotati di connettività di rete, in grado di raccogliere dati ambientali tramite sensori, condividerli con altri dispositivi via Internet e analizzarli (o intervenire) in modo automatico, senza bisogno dell'intervento umano. (Dwivedi *et al.*, 2021)

Per garantire questa comunicazione, i dispositivi IoT utilizzati per il *sensing*, l'azione e il monitoraggio sono connessi a Internet tramite tecnologie di comunicazione come, ad esempio, Wi-Fi, Bluetooth Low Energy (BLE), ZigBee e Near Field Communication (NFC). (Choudhary *et al.*, 2024) Il collegamento in rete di tutti gli elementi della filiera consente la raccolta e trasmissione continua di dati a server ad alta capacità di calcolo, dove l'elaborazione consente di prendere decisioni che possono essere immediatamente eseguite da altri dispositivi IoT. (Mistry *et al.*, 2020)

Nell'industria vitivinicola, l'integrazione di queste tecnologie ha migliorato significativamente la tracciabilità lungo l'intera filiera produttiva, grazie ai dati raccolti da sensori, tag RFID (*Radio Frequency Identification*) e *smart labels*, che monitorano e trasmettono in tempo reale informazioni sulle condizioni dei vigneti, sui processi di vinificazione e sulla distribuzione finale. (Izquierdo-Bueno *et al.*, 2024)

Ad esempio, per tracciare la qualità e l'integrità della biomassa vitivinicola dalla raccolta alla cantina, i tag RFID applicati ai contenitori delle uve si sono rivelati estremamente utili, permettendo di monitorare in tempo reale non solo la posizione, ma anche le condizioni di trasporto e conservazione del prodotto, elementi fondamentali per garantire la freschezza e l'integrità delle uve durante tutta la filiera. (Expósito *et al.*, 2013)

Parallelamente, l'utilizzo di sensori IoT per il monitoraggio costante di parametri critici quali temperatura e umidità, i cui dati vengono analizzati in tempo reale da sistemi di intelligenza artificiale, consente di mantenere condizioni ottimali durante l'affinamento del vino, prevenendo fenomeni di deterioramento come ossidazione e crescita microbica indesiderata, e assicurando così la qualità costante del prodotto finale. (Verzeletti *et al.*, 2023)

Un esempio significativo dell'applicazione dell'IoT in viticoltura è rappresentato dallo sviluppo di sistemi autonomi basati su sensori spettrali, progettati per misurare direttamente la riflettanza dei grappoli d'uva e monitorare *in situ* parametri chimici come i solidi solubili totali durante le diverse fasi di maturazione; tali dispositivi, concepiti come nodi terminali di reti di sensori wireless, integrano funzionalità di acquisizione dati, connettività e analisi automatizzata, consentendo un monitoraggio continuo e non distruttivo della qualità dell'uva dalla fase di invaiatura fino alla vendemmia. (Oliviera *et al.*, 2024)

In prospettiva, l'evoluzione delle tecnologie IoT, integrate con sistemi di intelligenza artificiale e sensori avanzati, potrà consentire un monitoraggio sempre più accurato e automatizzato non solo dei parametri qualitativi del vino, ma anche della gestione e valorizzazione della biomassa vitivinicola lungo l'intera filiera produttiva, favorendo processi più efficienti, tracciabili e sostenibili. (Ferro *et al.*, 2023; Adamashvili *et al.*, 2024)

2. SCOPO DEL LAVORO

Lo scopo di questo lavoro, svolto presso l'azienda Vegea Srl nell'ambito di un progetto dedicato alla valorizzazione degli scarti vitivinicoli, considerando che in Europa vengono prodotti annualmente circa 14,5 milioni di tonnellate di sottoprodotti dell'uva, è lo sviluppo di un dispositivo per il monitoraggio della vinaccia, concepito per selezionare la biomassa di qualità superiore da destinare a diverse applicazioni. Il progetto mira alla realizzazione di una tecnologia avanzata che integri sensori, Intelligenza Artificiale e Internet delle Cose (IoT), rendendo più efficiente il processo di caratterizzazione e gestione della vinaccia. Il progetto è sviluppato in collaborazione con un'azienda partner e un'università [nomi omessi per motivi di *privacy*].

Durante il tirocinio, le attività svolte hanno incluso un'approfondita ricerca bibliografica per individuare i fattori da monitorare indicativi dell'avanzamento della *shelf life* della vinaccia e per selezionare i parametri più rilevanti, fornendo un supporto scientifico agli ingegneri impegnati nello sviluppo del software e dell'hardware del dispositivo. Successivamente, è stata condotta un'analisi preliminare su modelli di regressione basati su tecniche di *machine learning* al fine di valutarne la potenziale integrazione nel sistema.

Inoltre, è stata studiata l'influenza delle antocianine sulla qualità e sulla *shelf life* della vinaccia tramite analisi delle correlazioni e applicazione dei modelli di *machine learning* già testati, valutandone anche la significatività statistica per identificare strategie in grado di ottimizzare le prestazioni del dispositivo e la selezione della biomassa.

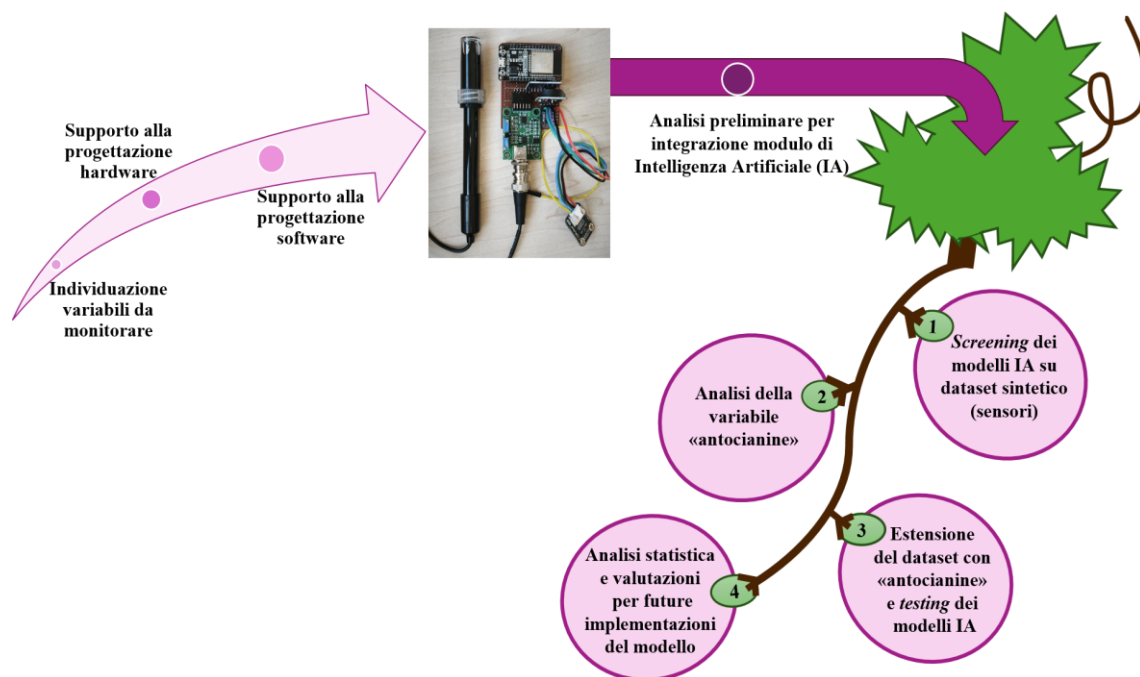


Figura 2.1. *Work-flow* del lavoro.

3. MATERIALI E METODI

3.2 COMPONENTE HARDWARE

3.2.1 MICROCONTROLLORE PRINCIPALE - ESP32-S3-WROOM-1

Il microcontrollore ESP32-S3-WROOM-1 (*Espressif Systems, Shanghai, Cina*) costituisce l'unità centrale di elaborazione del sistema. Esso integra un processore dual-core Tensilica LX7 con frequenza operativa massima di 240 MHz e supporto alle istruzioni vettoriali per applicazioni di intelligenza artificiale. Il dispositivo è dotato di connettività Wi-Fi 802.11 b/g/n e Bluetooth 5.0 LE, consentendo la comunicazione wireless con reti locali o dispositivi remoti. Il modulo include un'ampia disponibilità di linee GPIO (*General Purpose Input/Output*) configurabili e compatibili con i principali protocolli di comunicazione digitale, I²C, SPI, UART, PWM, oltre a convertitori analogico/digitali (ADC) a 12 bit per l'acquisizione diretta di segnali analogici da sensori esterni. Tali caratteristiche rendono il microcontrollore idoneo a sistemi *embedded* complessi e applicazioni IoT che richiedono capacità di elaborazione distribuita, gestione di periferiche eterogenee e connettività integrata.

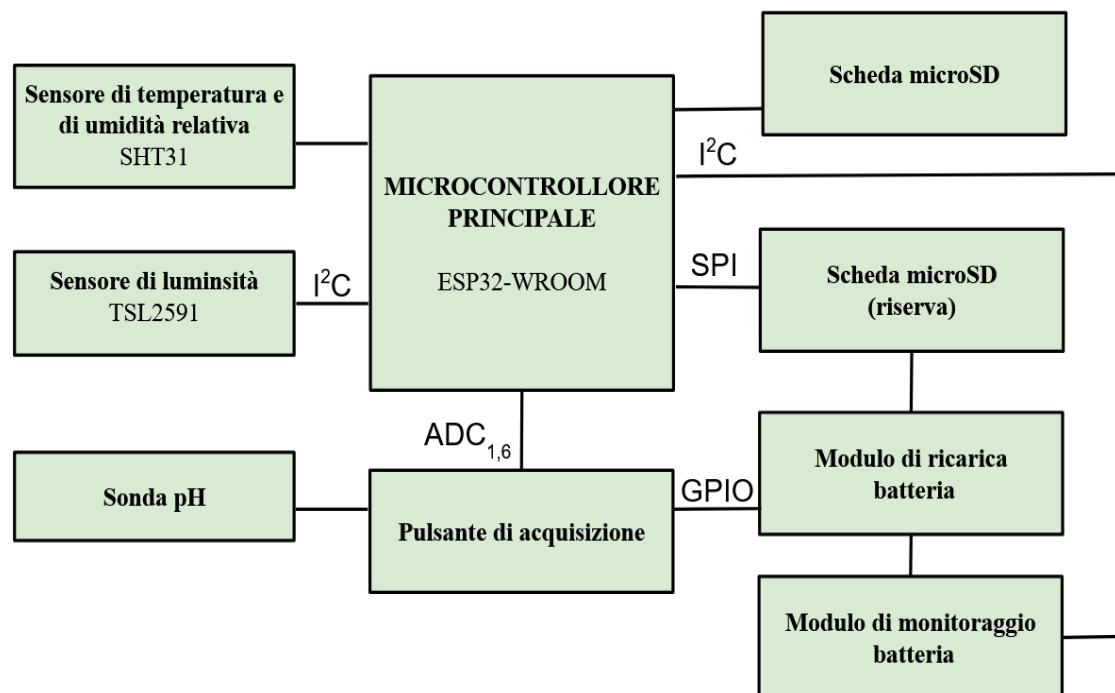


Figura 3.1 Architettura hardware del dispositivo di monitoraggio progettato.

3.1.2 SENSORE DI TEMPERATURA E UMIDITÀ - SHT31

Il sensore SHT31 (*Sensirion AG, Stäfa, Svizzera*), versione 1.3, è un trasduttore digitale per la misura combinata di temperatura e umidità relativa. Il dispositivo appartiene alla famiglia CMOSens® di Sensirion e integra, su un singolo circuito, l'elemento sensibile, il sistema di condizionamento del segnale, il convertitore analogico-digitale e l'interfaccia di comunicazione I²C. Il componente è progettato per operare con una tensione di alimentazione compresa tra 1.08 V e 3.6 V e supporta frequenze di comunicazione fino a 1 MHz. Le misure vengono fornite in formato digitale a 16 bit di risoluzione, con una precisione tipica di ± 0.3 °C per la temperatura e ± 1.8 % RH per l'umidità relativa. Il tempo di risposta, inferiore a 8 s, consente l'acquisizione affidabile anche in presenza di rapide variazioni delle condizioni termoigrometriche. Il sensore integra inoltre meccanismi di calibrazione e compensazione automatica che garantiscono stabilità metrologica e ridotta deriva a lungo termine, rendendolo idoneo per applicazioni di monitoraggio ambientale in sistemi embedded e dispositivi IoT.

3.1.3 REAL TIME CLOCK - DS3231

Il DS3231 (*Analog Devices Inc., Wilmington, Massachusetts, USA*) è il modulo responsabile della gestione temporale all'interno del dispositivo. Nello specifico, è un Real-Time Clock (RTC) ad alta precisione, progettato per fornire una misura accurata del tempo anche in assenza di alimentazione principale. È caratterizzato dall'integrazione di un oscillatore a compensazione di temperatura (TCXO), che riduce notevolmente gli errori di deriva tipici degli orologi digitali convenzionali. Il DS3231 ha una precisione di circa ± 2 ppm, corrispondente a uno scarto inferiore a 1 minuto all'anno. Inoltre, è dotato di un sistema di alimentazione secondaria a batteria tampone, che gli consente di continuare a funzionare anche quando il resto del dispositivo è spento o disconnesso dalla rete elettrica, garantendo così continuità nella misura del tempo.

3.1.4 SENSORE DI LUMINOSITÀ – TSL2591

Il TSL2591 (*AMS-OSRAM AG, Premstätten, Austria*) è un sensore fotometrico digitale ad alta dinamica, progettato per misurare la luminosità ambientale in un intervallo estremamente ampio, compreso tra 0,001 e oltre 100.000 lux. È dotato di due fotodiodi interni: uno a banda larga, sensibile sia alla luce visibile che a quella infrarossa, e uno infrarosso dedicato (IR). La

combinazione delle due misure consente di stimare accuratamente l'intensità luminosa percepita dall'occhio umano, compensando l'influenza della radiazione IR. Il dispositivo comunica con il microcontrollore ESP32 tramite interfaccia I²C e viene gestito dal firmware mediante la libreria *Adafruit_TSL2591*.

Una logica di *autorange* integrata nel firmware permette l'adattamento automatico del guadagno e del tempo di integrazione in funzione delle condizioni di illuminazione:

- se il livello di luce rilevato è basso, il guadagno viene incrementato fino a un tempo di integrazione di 600 ms, al fine di aumentare la sensibilità;
- in presenza di luce intensa, il guadagno viene ridotto fino a 100 ms per evitare fenomeni di saturazione del sensore.

Questa gestione dinamica consente di ottenere misure lineari e stabili in un'ampia gamma di condizioni ambientali, rendendo il sistema adatto sia per applicazioni indoor che per misurazioni in campo aperto.

3.1.5 SENSORE DI pH

Sonda analogica per il rilevamento del pH, collegata all'ingresso ADC1_6 (pin 36) del microcontrollore ESP32. È implementato nel software gestionale del hardware (il firmware) un algoritmo di filtraggio *multi-step* volto a garantire la stabilità e l'affidabilità della misura per evitare segnale soggetto a disturbi elettrici e variazioni lente.

L'elaborazione del segnale avviene secondo le seguenti fasi principali:

- **Acquisizione multipla:** vengono prelevati nove campioni consecutivi dal convertitore ADC a 12 bit (attenuazione 11 dB), in modo da coprire l'intervallo di tensione compreso tra 0 e 3,3 V.
- **Filtraggio per mediana:** i campioni vengono ordinati e ne viene estratto il valore mediano, riducendo l'influenza di eventuali picchi di rumore.
- **Scarto statistico:** i valori che si discostano dalla mediana di oltre ± 150 LSB vengono scartati; la media dei campioni rimanenti (*trimmed mean*) rappresenta il valore filtrato.
- **Conversione in tensione:** il valore medio grezzo viene convertito in volt tramite la funzione `esp_adc_cal_raw_to_voltage()`, basata sulla caratterizzazione reale dell'ADC (`esp_adc_cal_characterize()`).
- **Filtraggio IIR (*Infinite Impulse Response*):** viene applicato un filtro digitale del primo ordine, descritto dalla relazione

$$y_k = a \cdot y_{k-1} + (1 - a) \cdot v_k$$

dove ($a = 0.85$) rappresenta il coefficiente di smorzamento. Questo filtro attenua le oscillazioni a breve termine mantenendo una buona reattività alle variazioni reali.

Infine, la tensione filtrata viene trasformata in valore di pH mediante una relazione lineare di calibrazione:

$$pH = m \cdot V + q$$

dove m è la pendenza (*slope*) e q l'intercetta (*intercept*) della curva di taratura. I coefficienti m e q sono calibrabili dall'utente e salvati in memoria non volatile (NVS) tramite la libreria *Preferences*. La calibrazione può essere aggiornata:

- via porta seriale, con il comando SET_PH <m> <q>;
- oppure via MQTT, mediante il payload JSON:

```
{ "set_ph": { "slope": -9.09, "intercept": 21.00 } }
```

Questa architettura consente di eseguire la calibrazione direttamente in campo, senza necessità di ricompilare o ricaricare il firmware, garantendo maggiore flessibilità e semplicità operativa.

3.1.6 MICROSD - MEMORIA ESTERNA

La scheda MicroSD rappresenta il supporto di memoria esterno utilizzato nel sistema per la registrazione locale dei dati raccolti dai sensori ambientali. La MicroSD si interfaccia con il microcontrollore tramite il protocollo SPI (*Serial Peripheral Interface*), che consente una comunicazione veloce e affidabile. Il dispositivo utilizza una scheda MicroSD formattata in FAT32, un *file system* ampiamente supportato e versatile.

3.1.7 MODULO DI RICARICA BATTERIA - TP4056

Il TP4056 (IC Texas Instruments, USA) è il modulo di ricarica utilizzato per gestire in sicurezza la batteria ricaricabile del dispositivo. Si tratta di un caricatore lineare progettato specificamente per batterie agli ioni di litio (Li-Ion) o ai polimeri di litio (Li-Po) a singola cella. La corrente di carica è programmabile fino a un massimo di 1 A, permettendo di adattare il tempo di ricarica in base alla capacità della batteria impiegata.

3.1.8 MODULO MONITORAGGIO BATTERIA - MAX17043

Il sistema progettato integra un modulo per il monitoraggio della batteria, basato sul chip MAX17043 (Analog Devices, USA), entrambi compatibili con l'interfaccia di comunicazione

I²C. In particolare, il MAX17043 è un *fuel gauge* digitale progettato per batterie a singola cella agli ioni di litio. Il suo compito principale è stimare in tempo reale la percentuale di carica residua della batteria, fornendo un'indicazione accurata dello stato di carica (*State of Charge*, SoC).

Questo modulo si basa su un algoritmo brevettato di tipo “ModelGauge”, che consente di evitare complesse misurazioni di corrente o calibrazioni esterne, semplificando così l'integrazione nel sistema.

3.1.9 FIRMWARE DI CONTROLLO DEL DISPOSITIVO DI MONITORAGGIO

Il firmware della centralina del dispositivo di monitoraggio è sviluppato interamente in C++ sfruttando l'ecosistema Arduino per ESP32, una piattaforma *open-source* che combina semplicità di sviluppo e flessibilità hardware.

L'architettura software adotta un approccio modulare e fortemente disaccoppiato, con sottosistemi logici indipendenti tra loro, progettati per favorire chiarezza del codice, manutenibilità e scalabilità. Il sottosistema di gestione sensori si occupa dell'acquisizione dei dati ambientali tramite bus I²C per i sensori SHT31 e TSL2591 e canale ADC per la sonda analogica di pH, con letture filtrate e validate per ridurre il rumore e aumentare l'affidabilità. Il sottosistema di rete gestisce la connessione Wi-Fi in modalità stazione, con riconnessione automatica e supporto a IP statici o dinamici, integrando un modulo MQTT per la pubblicazione dei dati e il ricevimento di comandi remoti. La gestione temporale è affidata a NTP (via internet) e al modulo RTC DS3231, garantendo *timestamp* coerenti anche in assenza di rete. La registrazione dei dati avviene su scheda SD in formato CSV tramite la libreria SD.h, offrendo un backup locale indipendente dalla rete, mentre le impostazioni persistenti relative a Wi-Fi, MQTT e calibrazione pH sono memorizzate nella memoria *flash* interna tramite la classe Preferences, rendendole disponibili dopo riavvii.

Il ciclo operativo del firmware ruota attorno alle funzioni *setup()*, che inizializza librerie, pin, connessione di rete, orologio e pubblica un messaggio di “boot” su MQTT, e *loop()*, che gestisce tutte le operazioni in maniera non bloccante basata su eventi e timer, garantendo reattività anche in caso di errori temporanei di rete o periferiche.

Il firmware si avvale di librerie *open-source* selezionate tra cui WiFi.h, PubSubClient.h, Adafruit_SHT31.h, Adafruit_TSL2591.h, RTCLib.h, SD.h, ArduinoJson.h, Preferences.h e mbedtls/base64.h, che supportano rispettivamente la connettività di rete, la gestione MQTT, i sensori ambientali, il modulo RTC, il logging su scheda SD, la serializzazione JSON, la

persistenza dei parametri e la decodifica base64 utilizzata nella configurazione sicura delle credenziali.

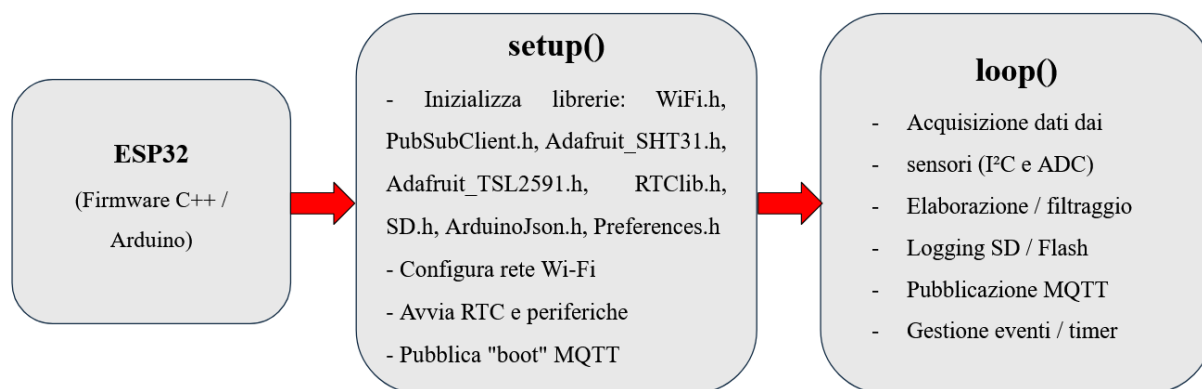


Figura 3.2 Architettura del Firmware del dispositivo di monitoraggio progettato.

Il firmware implementa meccanismi di riconnessione Wi-Fi e MQTT con backoff prudente, bufferizzazione minima per pubblicazione dei dati (le misure vengono inviate se rete e broker sono disponibili, altrimenti registrate su SD e seriale senza bloccare la logica), dimensionamento adeguato del payload MQTT (fino a 1024 B per JSON completi) e controlli di integrità sui pacchetti con range checking dei parametri. Sul piano della sicurezza, le credenziali possono essere fornite in base64 o URL-*encoded* e salvate in NVS, evitando l'hard-coding sensibile, mentre l'integrazione con TLS/MQTTS è possibile con broker compatibili.

Il firmware include strategie di resilienza, come riconnessione Wi-Fi/MQTT automatica, bufferizzazione minima con logging su SD se rete non disponibile, dimensionamento payload a 1024 B e validazioni sui parametri. La combinazione di CSV su seriale e JSON su MQTT, insieme a marcatura temporale coerente (ISO-8601 + epoch in CSV), permette fusioni dati affidabili, audit trail e integrazione con strumenti eterogenei (Excel, Python, pipeline cloud), preparando il terreno alla parte software di backend e frontend.

3.1.10 COMUNICAZIONE ARDUINO E SOFTWARE DI SUPERVISIONE

La comunicazione tra la centralina ESP32 del dispositivo e il software di supervisione si basa su due canali complementari: la porta seriale USB/UART, per il controllo locale e il debug, e il protocollo MQTT su rete Wi-Fi, per la telemetria remota e il telecontrollo. La porta seriale consente l'emissione continua di righe CSV con i valori misurati, facilmente interpretabili da software desktop o strumenti come terminali seriali, Python o MATLAB, e permette l'invio di comandi testuali per configurazione rapida, diagnostica e test.

Un esempio di riga CSV di telemetria è il seguente:

```
epoch,tempC,humRH,lux,ph
1728561005,24.37,45.12,341.55,6.82
```

dove epoch rappresenta l'Unix time (secondi dal 1970) calcolato tramite NTP o RTC DS3231, mentre "tempC", "humRH", "lux" e "ph" riportano rispettivamente temperatura in gradi Celsius, umidità relativa in %, illuminamento in lux e valore di pH. Il firmware gestisce comandi come MEASURE per misurazione immediata, SYNC per sincronizzazione dell'orologio, HEATER ON|OFF per il riscaldatore SHT31, SET_PERIOD per il periodo di invio, SET_PH per la taratura del sensore pH e comandi per configurazione Wi-Fi e MQTT. Il *parsing* dei comandi è robusto e genera linee di stato interpretabili in modo deterministico, ad esempio:

```
status=set_wifi;result=ok
wifi=192.168.1.50;mqtt=connected
```

Il canale MQTT abilita scenari IoT distribuiti, con invio asincrono di dati verso un broker (*on-premise* o cloud) e integrazione con *dashboard* e servizi di analisi, consentendo la pubblicazione delle misure su `nomedispositivo/env`, dello stato del dispositivo su `nomedispositivo/status` e la sottoscrizione ai comandi su `nomedispositivo/cmd`. I campi numerici possono essere nulli in caso di misure non disponibili (es. sensore assente o saturazione), mantenendo coerenza semantica del payload. La *timestamp* è in ISO-8601 con offset di fuso, derivata da NTP o, in fallback, da RTC.

Un esempio di payload JSON per la telemetria è:

```
{
  "ts": "2025-10-10T12:34:56+02:00",
  "id": "NOMEDISPOSITIVO-001",
  "tempC": 24.37,
  "humRH": 45.12,
  "lux": 341.55,
  "ph": 6.82,
  "src": "timer"
}
```

mentre il *payload* relativo allo stato del dispositivo è:

```
{
```

```
"ts": "2025-10-10T12:35:00+02:00",
"id": "NOMEDISPOSITIVO-001",
"fw": "nomedispositivo-esp32 v2.0",
"status": "connected",
"wifi": "192.168.1.50",
"rtc": "ok"
}
```

I comandi MQTT sono anch'essi strutturati in JSON e, ad ogni esecuzione, producono un riscontro sul topic di stato (es. esito *_ok, *_invalid), assicurando la tracciabilità delle operazioni e la coerenza dei dati raccolti in contesti distribuiti.

Comandi via MQTT (nomedispositivo/cmd) – JSON ad azioni atomiche:

```
{ "cmd": "measure" }           // misura immediata
{ "cmd": "sync" }             // sync NTP/RTC
{ "cmd": "get_info" }        // pubblica stato
{ "heater": "on" }           // heater SHT31 on/off
{ "set_period": 30 }         // periodo invio (s) 5..3600
{ "set_ph": { "slope": -9.09, "intercept": 21.0 } }
```

3.2 COMPONENTE SOFTWARE

3.2.1 BACKEND

Il backend costituisce il nucleo logico dell'intero sistema software, rappresentando il livello centrale di orchestrazione tra le diverse componenti dell'applicazione: l'interfaccia utente, il database, i dispositivi IoT e il modulo di intelligenza artificiale.

È stato sviluppato utilizzando Laravel, un *framework open-source* basato su PHP 8+, linguaggio di programmazione *server-side* ampiamente utilizzato per la realizzazione di applicazioni web dinamiche. L'indicazione "8+" fa riferimento all'adozione della versione 8 o successive, che introducono funzionalità moderne e prestazioni ottimizzate.

A supporto del backend è stato integrato MySQL, un sistema di gestione di database relazionali *open-source*, scelto per la sua stabilità, scalabilità e ampia diffusione nel panorama applicativo.

Nel complesso, il backend è responsabile della gestione dei dati, dell'applicazione della logica di *business* e della comunicazione tra i diversi attori del sistema, fungendo da ponte tra i componenti fisici (dispositivi e sensori) e i servizi software (interfaccia utente).

Il backend implementa una serie di funzionalità fondamentali per il corretto funzionamento dell'intero sistema, riassunte di seguito.

- Gestione utenti e ruoli: il sistema consente di definire e gestire utenti con ruoli differenziati in base al loro ruolo nella filiera produttiva (es. ricercatori, tecnici, operatori). A ciascun ruolo sono associati permessi specifici, che garantiscono un accesso controllato alle risorse e alle funzionalità, migliorando sia la sicurezza che l'usabilità della piattaforma.
- Gestione dei lotti e della *supply chain*: tramite operazioni CRUD (*Create, Read, Update, Delete*) è possibile registrare nuovi lotti di vinaccia, aggiornarne lo stato e monitorarne la tracciabilità lungo tutta la catena produttiva. Questo modulo supporta una gestione efficiente e trasparente dei dati relativi ai processi di trasformazione e valorizzazione dei sottoprodotti.
- Esposizione di API RESTful: il *backend* fornisce un insieme di API (*Application Programming Interface*) conformi allo stile architetturale RESTful. Queste API, basate su URL (*Uniform Resource Locator*), semantiche e operazioni standard (GET, POST, PUT, DELETE), consentono una comunicazione bidirezionale con:
 - il *frontend*, per la visualizzazione, modifica e gestione dei dati da parte dell'utente;
 - i dispositivi IoT, da cui vengono ricevuti i dati ambientali rilevati dai sensori;
 - un eventuale modulo di intelligenza artificiale, previsto per future implementazioni, a cui potranno essere inviate richieste di previsione della *shelf life* della vinaccia.

Per quanto riguarda sicurezza e tracciabilità, l'autenticazione avviene tramite token, chiavi univoche assegnate agli utenti che permettono di identificarli nelle richieste successive senza dover effettuare nuovamente il login. Le API sono protette in modo che solo utenti autorizzati possano accedervi o modificarne le risorse. Inoltre, un sistema di logging registra in modo persistente tutte le attività rilevanti (accessi, modifiche, errori, ecc.), garantendo così trasparenza operativa e la possibilità di effettuare audit o analisi successive.

Nel complesso, il backend rappresenta l'infrastruttura logica di riferimento dell'intero ecosistema IoT, fungendo da mediatore tra i dispositivi fisici e i servizi software. La sua architettura modulare e scalabile garantisce affidabilità, sicurezza e tracciabilità delle

operazioni, assicurando al contempo un'integrazione efficiente con i moduli di intelligenza artificiale e con l'interfaccia utente per la gestione e l'analisi dei dati.

3.2.2 FRONTEND

L'interfaccia utente è progettata per essere intuitiva, funzionale e adattabile ai diversi ruoli presenti nella filiera produttiva, garantendo un'esperienza d'uso efficiente e personalizzata per tutti gli utenti, inclusi quelli senza competenze tecniche specifiche, supportando così decisioni informate e tempestive.

La tecnologia utilizzata per lo sviluppo del frontend è Angular (versione 14 o superiore), un *framework* sviluppato da Google scritto in TypeScript, un linguaggio costruito sopra JavaScript che ne migliora la robustezza e la manutenzione del codice.

Tra le funzionalità principali troviamo:

- *Dashboard* interattiva: una pagina riassuntiva che mostra indicatori chiave, grafici e notifiche, offrendo una panoramica immediata dello stato dei lotti, delle condizioni ambientali e delle eventuali previsioni generate dal modulo di intelligenza artificiale.
- Visualizzazione dei dati storici: grafici dinamici, timeline e strumenti di confronto tra i diversi lotti, che permettono un'analisi dettagliata e rapida dell'evoluzione delle variabili ambientali e dei risultati.
- Sezione dedicata all'Intelligenza artificiale (per implementazione futura): visualizza i risultati delle previsioni sulla *shelf life* e l'andamento stimato della degradazione della vinaccia.
- Gestione della *supply chain*: consente di tracciare il percorso di ogni lotto, verificando i dati raccolti in ogni fase del processo produttivo per garantire trasparenza e controllo.
- Interfaccia responsive: progettata per adattarsi perfettamente a dispositivi desktop e mobile, assicurando accessibilità e usabilità in qualsiasi contesto, anche sul campo.

In sintesi, il *frontend* costituisce il punto di interazione tra l'utente e il sistema, traducendo la complessità dei processi sottostanti in un'interfaccia accessibile, chiara e funzionale. La progettazione basata su Angular e l'approccio responsive assicurano un'esperienza d'uso coerente e fruibile su qualsiasi dispositivo, consentendo un monitoraggio efficace e in tempo reale dei dati acquisiti e, in futuro, anche le previsioni di *shelf life*.

3.2.4 GESTIONE E RICEZIONE DEI DATI NEL SOFTWARE DI SUPERVISIONE

I sensori fisici del sistema IoT progettato possono trasmettere le misurazioni ambientali al sistema informatico seguendo due modalità operative distinte, determinate dal protocollo di comunicazione adottato: HTTP REST API o MQTT.

L'approccio basato su HTTP REST API utilizza HTTP, il protocollo standard per la comunicazione su Internet, in combinazione con interfacce REST (*Representational State Transfer*), che permettono ai sensori di inviare dati tramite richieste strutturate, come ad esempio una POST verso l'endpoint /invia-lettura. L'approccio MQTT (*Message Queuing Telemetry Transport*), invece, si basa su un protocollo di messaggistica leggero e ottimizzato per dispositivi a basso consumo energetico o con banda limitata. È particolarmente adatto in contesti IoT ad alta frequenza di trasmissione, grazie alla sua efficienza e alla capacità di gestire numerosi sensori attivi in tempo reale. In questo scenario, i sensori pubblicano i dati su topic specifici (ad esempio lotto/123/temperatura), mentre il server si sottoscrive ai topic di interesse per ricevere automaticamente i messaggi non appena vengono emessi.

La motivazione della scelta di una struttura ibrida che preveda la presenza simultanea di entrambi i protocolli offre maggiore flessibilità e resilienza al sistema. In particolare:

- Se un sensore non supporta MQTT, può comunque trasmettere i dati tramite HTTP
- In caso di malfunzionamento del broker MQTT, è possibile utilizzare HTTP come canale di fallback
- MQTT è ideale per l'invio continuo di dati a bassa latenza (es. ogni 30 secondi), mentre HTTP si presta meglio per operazioni di configurazione, invii sporadici o interazione con sistemi esterni (es. IA e backend)

In sintesi, MQTT è il canale privilegiato per la raccolta automatica e in tempo reale dei dati ambientali, mentre HTTP REST API rappresenta il collante universale che consente l'interazione tra sensori, backend e modulo di intelligenza artificiale.

Il sistema è progettato per garantire un'acquisizione completamente automatizzata delle misurazioni ambientali, riducendo al minimo l'intervento umano e aumentando l'affidabilità del processo. Ogni lettura proveniente dai sensori viene associata in modo univoco a un lotto specifico di vinaccia, permettendo di collegare in maniera precisa i dati raccolti al relativo processo produttivo. Inoltre, ogni dato è marcato con un *timestamp*, cioè con la data e l'orario esatto della rilevazione, assicurando una tracciabilità temporale dettagliata.

Una volta ricevute, le informazioni vengono archiviate automaticamente all'interno del database relazionale, senza necessità di operazioni manuali, creando una base solida per l'analisi successiva, sia da parte del modulo di intelligenza artificiale sia da parte degli operatori umani tramite l'interfaccia utente.

Questa architettura garantisce una tracciabilità completa lungo tutta la filiera, dalla raccolta del dato ambientale fino alla visualizzazione della previsione sulla *shelf life*. Inoltre, consente di ottenere previsioni affidabili grazie alla disponibilità di dati storici accurati e costantemente aggiornati. Infine, l'intero sistema è progettato per operare in modo robusto anche in ambienti remoti o con connettività intermittente, assicurando continuità operativa e mantenimento delle funzionalità critiche anche in condizioni non ideali.

3.3 MODULO DI INTELLIGENZA ARTIFICIALE

3.3.1 AMBIENTE DI SVILUPPO E STRUMENTI SOFTWARE

L'intero *workflow* di sviluppo dei modelli di Intelligenza Artificiale è stato realizzato in Python (versione 3.13.7), all'interno dell'ambiente di programmazione Visual Studio Code, scelto per la sua flessibilità e per l'integrazione con i principali strumenti di *data science*.

Per l'elaborazione e la gestione dei dati sono state impiegate le librerie Pandas e NumPy, quest'ultima utilizzata per le operazioni numeriche e la manipolazione di *array* multidimensionali. La visualizzazione dei risultati e delle metriche di confronto è stata realizzata mediante Matplotlib e Seaborn, che hanno permesso di produrre grafici comparativi e rappresentazioni dettagliate delle *performance* dei modelli.

L'addestramento e la valutazione dei modelli di regressione sono stati effettuati tramite scikit-learn e TensorFlow/Keras (versione 2.20). Scikit-learn è stata impiegata per l'implementazione dei modelli classici di regressione, tra cui *Regressione Lineare multipla* (MLR), *Random Forest Regressor* (RFR), *Gradient Boosting Regressor* (GBR), *Multilayer Perceptron Regressor* (MLP) e *Support Vector Regression* (SVR).

I dati sono stati suddivisi in *training set* (80%) e *test set* (20%) per garantire una valutazione indipendente delle prestazioni e prevenire fenomeni di *overfitting*. Per i modelli di rete neurale profonda, in particolare la rete *fully connected* e la Improved-ResNet sviluppata *ad hoc*, è stato utilizzato TensorFlow 2.20 con l'API Keras, che consente una gestione modulare dell'architettura e del processo di addestramento.

I dati di input sono stati normalizzati mediante StandardScaler e MinMaxScaler, strumenti forniti da scikit-learn, al fine di garantire la corretta scalatura delle feature numeriche e ottimizzare la convergenza dei modelli. Tutte le simulazioni sono state eseguite su macchina locale, senza l'impiego di GPU esterne o ambienti di calcolo distribuito.

3.3.2 GENERAZIONE DEI DATASET SINTETICI

Per la sperimentazione virtuale, sono stati generati due dataset simulati della vinaccia, rispettando le caratteristiche dei sensori presenti nel sistema IoT per quanto concerne temperatura, umidità relativa, luminosità e pH. Il primo dataset, *dataset A*, comprende solo le variabili ambientali e la *shelf life* della vinaccia, mentre il secondo, *dataset B*, include, in aggiunta, il livello di antocianine, consentendo di effettuare un'analisi esplorativa dei dati (EDA) e valutare la sua possibile integrazione nel dispositivo IoT. Tra i due dataset, l'unica differenza risiede quindi nella presenza della variabile *anthocyanin_level* nel secondo.

I dataset sono stati generati con 1.000 campioni ciascuno, simulando rilevazioni a intervalli di 30 minuti a partire dal 1 ottobre 2025, assunto come data di riferimento ipotetica. La scelta del *timestep* è coerente con la frequenza di acquisizione prevista dal dispositivo reale, al fine di riprodurre in modo quanto più verosimile possibile il comportamento del sistema IoT. La generazione dei dati è stata effettuata con fissazione del generatore casuale, così da garantire la riproducibilità della simulazione.

Le variabili ambientali sono state simulate come segue: la temperatura con rumore di $\pm 0,3$ °C, l'umidità relativa con ± 2 % di accuratezza, la luminosità log-normale tra 0,001 e 100.000 lux con rumore ± 5 %, e il pH filtrato per escludere misure fuori tolleranza. Il livello di antocianine è stato generato come funzione delle condizioni ambientali, diminuito in presenza di stress termici, luminosi o di umidità e modulato dal pH.

Per ogni campione, il codice Python calcolava direttamente anche la *shelf life*, modellata come combinazione lineare di variabili ambientali e chimiche, in linea con approcci di regressione lineare multivariata riportati in letteratura (Torres-Sánchez *et al.*, 2020). La stessa formulazione empirica è stata applicata a entrambi i dataset sintetici, includendo il termine *anthocyanin_level* esclusivamente nel dataset B e aggiungendo un termine di rumore casuale ϵ .

In particolare, la *shelf life* è stata calcolata secondo la seguente relazione:

$$\text{shelf_life_days} = 50 + 0.8 \cdot \text{anthocyanin_level} - 0.9 \cdot (\text{tempC} - 25) - 0.5 \cdot (\text{humRH} - 55) - 0.001 \cdot \text{lux} - 2.5 \cdot (\text{ph} - 6.5) + \epsilon,$$

dove ϵ rappresenta il rumore casuale normalmente distribuito per simulare variazioni sperimentali.

I dati sono stati normalizzati e salvati in formato CSV, pronti per l'analisi esplorativa e per l'addestramento dei modelli di intelligenza artificiale sviluppati sempre in Python.

La **Figura 3.3** mostra un esempio della struttura dei *dataset A* e *B*, organizzati su base temporale mediante timestamp e differenziati unicamente dalla presenza della variabile *anthocyanin_level* nel *dataset B*.

<i>Dataset A</i>	Campione n°	epoch,tempC,humRH,lux,ph,shelf_life_days
	1	1761107400.00,27.30,39.46,93912.98,6.25,90.42
	2	1759658400.00, 25.67, 83.85, 23993.29,6.22, 81.09

<i>Dataset B</i>	Campione n°	epoch,tempC,humRH,lux,ph,anthocyanin_level, shelf_life_days
	1	1761107400.00,24.30,69.46,93912.98,6.25,76.80,89.40
	2	1759392000.00, 21.79, 53.32, 17243.17, 6.92, 76.82, 97.18

Figura 3.3. Esempio della struttura dei dataset sintetici (*A* e *B*), organizzati su base temporale mediante *timestamp*; le variabili ambientali e chimiche sono visualizzate con colori differenti, con inclusione della variabile *anthocyanin_level* nel *dataset B*.

3.3.3 ANALISI ESPLORATIVA DEI DATI SUL LIVELLO DI ANTOCIANINE

Per esplorare le relazioni tra i livelli di antocianine nella vinaccia e la *shelf life*, al fine di valutare la possibilità di implementare tale parametro nel dispositivo IoT già costruito, il secondo dataset generato è stato analizzato mediante strumenti di analisi esplorativa dei dati (EDA), utilizzando lo stesso ambiente di sviluppo dei modelli di Intelligenza Artificiale. Sono state calcolate statistiche descrittive per tutte le variabili numeriche e visualizzate le distribuzioni tramite istogrammi. La correlazione tra le variabili è stata valutata mediante una *heatmap*, mentre la relazione specifica tra livello di antocianine e *shelf life* è stata analizzata tramite scatter plot e calcolo del coefficiente di correlazione di Pearson. Infine, è stato generato un pairplot per le principali variabili ambientali e di composizione chimica (temperatura, umidità relativa, illuminazione, pH, livello di antocianine e *shelf life*, al fine di identificare pattern, relazioni lineari e non lineari e potenziali *outlier*.

3.3.4 MODELLI DI INTELLIGENZA ARTIFICIALE UTILIZZATI

3.3.4.1 REGRESSIONE LINEARE MULTIPLA

Nel presente lavoro, il modello di regressione lineare, importata da `sklearn.linear_model`, è stata utilizzata come modello di base per la predizione della *shelf life* della vinaccia. I dati sono stati standardizzati tramite `StandardScaler` per garantire una corretta scala delle *features*. Questo modello assume una relazione lineare tra le variabili ambientali (temperatura, umidità, illuminazione, pH) e la *shelf life*, fornendo una *baseline* interpretativa semplice e immediata. Non sono stati applicati parametri aggiuntivi oltre ai valori di *default*.

3.3.4.2 RANDOM FOREST REGRESSOR (RFR)

Dal punto di vista implementativo, il *Random Forest Regressor* è stato importato da `sklearn.ensemble` e configurato con 300 alberi decisionali (`n_estimators=300`) e `random_state=42` per garantire la riproducibilità. I dati sono stati standardizzati mediante `StandardScaler` prima dell'addestramento.

3.3.4.3 GRADIENT BOOSTING REGRESSOR (GBR)

Il *Gradient Boosting Regressor*, anch'esso importato da `sklearn.ensemble`, è stato configurato con 300 alberi e *learning rate* pari a 0.05, con `random_state=42`. I dati standardizzati sono stati utilizzati per l'addestramento. Questo approccio addestra i modelli in sequenza, correggendo progressivamente gli errori dei modelli precedenti, migliorando così le prestazioni complessive su dataset con relazioni non lineari e pattern complessi.

3.3.4.4 SUPPORT VECTOR REGRESSION (SVR)

La *Support Vector Regression*, importata da `sklearn.svm`, è stata configurata con kernel RBF (*Radial Basis Fuction*), parametro `C=50` ed `epsilon=0.2`, e addestrata sui dati standardizzati. Questo modello è in grado di gestire relazioni non lineari e limitare l'impatto di valori anomali, ottimizzando la funzione obiettivo per mantenere gli errori entro un certo margine, risultando robusto su dataset con variabilità moderata.

3.3.4.5 MULTILAYER PERCEPTRON REGRESSOR IN SCIKIT-LEARN

La rete neurale *Multilayer Perceptron Regressor*, importata da `sklearn.neural_network`, è stata implementata con tre *layer* nascosti di 64, 32 e 16 neuroni, funzione di attivazione ReLU e `max_iter=1000`. I dati sono stati standardizzati tramite `StandardScaler`. Questo modello consente di catturare relazioni non lineari tra le variabili e di apprendere pattern complessi nei dati, rappresentando un primo passo verso modelli di deep learning più sofisticati.

3.3.4.6 MULTILAYER PERCEPTRON REGRESSOR IN KERAS

Un modello di rete neurale *fully connected* (Multilayer Perceptron, MLP) è stato sviluppato utilizzando la piattaforma TensorFlow/Keras (v.2.20), riconosciuta come un framework robusto per l'implementazione di soluzioni di *deep learning* in diversi ambiti applicativi. (Pang *et al.*, 2019; Rahim *et al.*, 2025) L'architettura del modello è composta da due *layer* nascosti contenenti rispettivamente 64 e 32 neuroni, ciascuno attivato tramite funzione ReLU, seguiti da un *layer* di output lineare per la stima di variabili continue. I dati in ingresso sono stati normalizzati mediante `MinMaxScaler` per uniformare la scala delle *features*. L'addestramento è stato condotto per 300 epoche con batch size di 16, adottando l'ottimizzatore Adam, la funzione di perdita MSE e la metrica MAE. Questa configurazione è stata scelta per favorire la convergenza del modello e minimizzare l'errore medio assoluto delle predizioni.

3.3.4.7 IMPROVED RESNET IN KERAS

Il modello ResNet migliorato è una rete neurale convoluzionale sviluppata *ad hoc* con TensorFlow/Keras, progettata per catturare pattern complessi tra le *features* ambientali e la *shelf life* della vinaccia. L'input, costituito dalle quattro *features*, viene rimodellato per essere compatibile con la convoluzione 1D selezionata in questa applicazione. La rete inizia con un *layer* convoluzionale di 32 filtri e kernel di dimensione 3, seguito da due *residual block*: il primo con 32 filtri, il secondo con 64. Ogni blocco residuo combina un *inception block* con convoluzioni di dimensione 1, 3 e 5, e un SE block (*Squeeze and Excitation block*) che modula l'importanza dei canali. Per adattare la *shortcut* al numero di filtri, viene applicata una convoluzione 1x1. Dopo i blocchi residui, un *layer* convoluzionale finale da 128 filtri precede un *global average pooling*, seguito da un *layer fully connected* con 64 neuroni e dropout 0.2, e infine il *layer* di output lineare. La rete è stata addestrata per 80 epoche, *batch size* 16, con

Adam e *learning rate* 1e-3, loss MSE e metriche MAE, usando dati standardizzati tramite StandardScaler. Questa architettura consente di modellare relazioni non lineari e di sfruttare informazioni multiscala tra le *features*, migliorando la capacità predittiva complessiva.

3.3.5 METRICHE DI VALUTAZIONE DELLE PRESTAZIONI DEI MODELLI

Le prestazioni dei modelli di regressione sono state valutate utilizzando tre metriche principali:

- **Mean Squared Error (MSE)**: misura la media dei quadrati degli errori tra i valori predetti (\hat{Y}_i) e quelli osservati (Y_i).

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Valori più bassi indicano una migliore accuratezza del modello. L'unità di misura corrisponde al quadrato dell'unità della variabile predetta (giorni², per la *shelf life*).

- **Mean Absolute Error (MAE)**: calcola la media degli errori assoluti tra predizioni e valori reali.

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|$$

Fornisce una misura intuitiva della deviazione media delle predizioni rispetto ai valori osservati. L'unità di misura è la stessa della variabile predetta (giorni, nel caso della *shelf life*).

- **Coefficiente di determinazione (R^2)**: quantifica la proporzione della varianza dei dati spiegata dal modello.

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

Dove (\bar{Y}) è la media dei valori osservati. R^2 è un valore adimensionale compreso tra $(-\infty)$ e 1; valori più vicini a 1 indicano un migliore adattamento ai dati.

Le metriche R^2 , MAE e MSE sono state scelte poiché forniscono una valutazione complementare delle prestazioni dei modelli: il coefficiente di determinazione (R^2) misura la capacità esplicativa globale, mentre MAE e MSE quantificano l'accuratezza delle previsioni in termini assoluti e quadratici. Questa combinazione consente di confrontare modelli con caratteristiche diverse (lineari, non lineari, neurali) in modo equilibrato, tenendo conto sia della precisione sia della robustezza delle stime.

In particolare, l'utilizzo congiunto di R^2 , MAE e MSE ha permesso di valutare in modo oggettivo le prestazioni dei modelli nella predizione della *shelf life* della vinaccia, fornendo indicazioni sia sulla capacità esplicativa complessiva (R^2) sia sull'entità media e quadratica degli errori (MAE, MSE).

Nell'identificazione dei modelli migliori, è stato considerato non solo un valore di R^2 elevato, ma anche un MAE basso, così da privilegiare soluzioni che garantissero un buon adattamento ai dati e al contempo una deviazione media degli errori contenuta, assicurando predizioni affidabili e stabili.

3.3.6 BOOTSTRAPPING

Il calcolo dei p-value per le differenze tra le metriche di prestazione dei modelli è stato effettuato mediante bootstrap, implementato in Python. Per ciascuna statistica di interesse $s(X)$ (ad esempio R^2 , MAE o MSE), sono stati generati N campioni bootstrap $X^{*1}, X^{*2}, \dots, X^{*N}$ dalla distribuzione sotto l'ipotesi nulla H_0 . Su ciascun campione bootstrap X^{*i} , è stato calcolato il valore della statistica $s(X^{*i})$. La distribuzione dei valori $\{s(X^{*1}), s(X^{*2}), \dots, s(X^{*N})\}$ fornisce una stima empirica della distribuzione nulla di $s(X)$, e il p-value bootstrap è stato calcolato come proporzione di campioni per cui $s(X^{*i}) \geq s(X)_{obs}$, con $s(X)_{obs}$ il valore osservato della statistica sul dataset originale:

$$p_{boot} = \frac{\text{numero di campioni bootstrap con } s(X^{*i}) \geq s(X)_{obs}}{N}$$

Questo approccio è stato applicato a tutti i 7 modelli testati, calcolando la differenza di R^2 , MAE e MSE tra i due dataset, per valutare se il miglioramento osservato nei modelli con il secondo dataset, che includeva una variabile aggiuntiva, fosse statisticamente significativo, utilizzando come soglia $\alpha=0.05$. Tale procedura permette di stimare la significatività delle differenze senza

assumere una distribuzione parametrica nota dei dati, garantendo robustezza anche con dataset limitati. (Vetura 2010)

In prima analisi è stato eseguito il bootstrap sul dataset B tale e quale. In seconda analisi è stata introdotta una *data augmentation* numerica per simulare la variabilità naturale delle misure di laboratorio, moltiplicando ogni valore numerico delle misurazioni ipotetiche delle *features* per un fattore casuale compreso tra 0.85–1.15 (rumore $\pm 15\%$) o 0.75–1.25 (rumore $\pm 25\%$), generando nuove configurazioni realistiche dei dati; in questa fase non è stato applicato alcun filtro sul pH, al fine di massimizzare la rappresentatività del campione.

Per ciascun livello di rumore sono stati quindi eseguiti esperimenti bootstrap ripetuti (da 300 a 1000 iterazioni), addestrando e validando i modelli in modo indipendente a ogni ciclo. Il modello Keras è stato testato fino a 1000 iterazioni, mantenendo sempre 100 epoch per ogni addestramento, in coerenza con quanto utilizzato durante la fase di test del modello. Gli altri modelli supervisionati (ad eccezione di ResNet) sono stati valutati con lo stesso numero di iterazioni.

4. RISULTATI E DISCUSSIONE

4.1 DISPOSITIVO IOT SVILUPPATO

Il sistema IoT sviluppato nell'ambito del progetto alla base della presente tesi, finalizzato al recupero delle bucce di uva rossa derivanti dalla vinificazione secondo i principi dell'economia circolare, ha dimostrato un funzionamento complessivamente stabile e affidabile, soddisfacendo gli obiettivi progettuali prefissati in termini di monitoraggio ambientale, gestione energetica e comunicazione dati. Le prove effettuate hanno evidenziato una corretta integrazione tra i moduli hardware e il firmware di controllo, con un'acquisizione regolare dei dati provenienti dai sensori e una trasmissione efficiente al backend remoto tramite protocollo MQTT. In condizioni di assenza di rete, la memorizzazione locale su scheda microSD ha garantito la continuità operativa del sistema e la successiva sincronizzazione automatica dei dati, confermando la resilienza dell'architettura proposta.

Dal punto di vista metrologico, i sensori impiegati hanno fornito misure accurate e consistenti. Il modulo SHT31 ha garantito una rilevazione precisa di temperatura e umidità relativa, con variazioni minime tra le acquisizioni successive e una buona risposta anche ai cambiamenti rapidi delle condizioni ambientali. Il sensore TSL2591, grazie all'ampia dinamica di misura e alla logica di *autorange*, ha mostrato un comportamento lineare sia in ambienti scarsamente illuminati che in condizioni di luce intensa. La sonda di pH, supportata dall'algoritmo di filtraggio e calibrazione implementato nel firmware, ha restituito dati stabili e ripetibili, riducendo sensibilmente l'influenza del rumore elettrico. Il DS3231, infine, ha mantenuto un'elevata precisione temporale, assicurando la corretta marcatura temporale dei dati anche durante le fasi di sospensione dell'alimentazione.

Sul piano energetico, il sistema ha mostrato un'efficiente gestione dell'alimentazione. Il modulo TP4056 ha garantito una ricarica stabile della batteria, con indicatori visivi di stato perfettamente funzionanti, mentre il circuito di monitoraggio basato su MAX17043 ha fornito letture affidabili dello stato di carica (SoC). L'integrazione di logiche di risparmio energetico nel firmware ha consentito di estendere l'autonomia operativa del dispositivo, rendendolo idoneo a misurazioni di lunga durata in ambienti non presidiati o privi di alimentazione continua.

Dal punto di vista software, il backend sviluppato in Laravel ha gestito in modo efficiente la ricezione, l'elaborazione e l'archiviazione dei dati ambientali, mentre il frontend in Angular ha offerto un'interfaccia utente chiara e reattiva, facilitando la visualizzazione dei dati, la

consultazione dei log e la gestione dei lotti. La stabilità del sistema è stata confermata anche da un utilizzo prolungato in ambiente di test, senza riscontrare anomalie significative o interruzioni del servizio. Tuttavia, per motivi di riservatezza e *privacy* aziendale, i dettagli quantitativi delle prove e dei test non possono essere pubblicati.

Complessivamente, il prototipo ha dimostrato di soddisfare i requisiti di affidabilità, precisione e scalabilità, risultando un sistema robusto e facilmente integrabile in scenari di monitoraggio agroindustriale. La progettazione modulare consente di estendere le funzionalità del dispositivo, ad esempio attraverso l'integrazione del modulo di intelligenza artificiale, senza modificare le componenti hardware o il firmware principale. Ciò rende il sistema predisposto per un'evoluzione futura verso un modello predittivo completo, in grado di stimare automaticamente la *shelf life* della vinaccia sulla base dei dati raccolti.

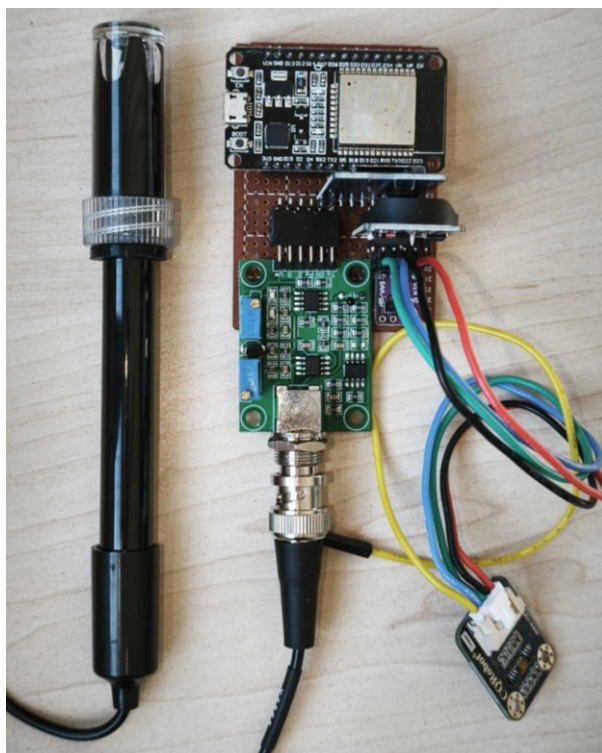


Figura 4.1 Prototipo del dispositivo progettato.

4.2 RISULTATI ANALISI COMPARATIVA DEI MODELLI DI IA TESTATI (DATASET A)

La selezione dei modelli di regressione è stata guidata dalla necessità di individuare un algoritmo predisposto a operare con la tipologia di dati che sarà generata dal sistema IoT, costituita da *timestamp* associati a misurazioni istantanee di temperatura, umidità relativa, luminosità e pH per ciascun campione. In tale contesto, sono stati scelti modelli in grado di

gestire relazioni non lineari, variazioni continue e un numero limitato di *features*, rappresentative delle condizioni ambientali monitorate dal dispositivo.

In questa prima fase di ricerca preliminare, è stato impiegato il primo dataset simulato, il *dataset A*, inizialmente composto da 1000 campioni, successivamente ridotto a 785 osservazioni valide in seguito all'applicazione dei filtri di qualità sul pH e dei limiti di plausibilità fisica relativi a temperatura, umidità e luminosità. Tale riduzione rispecchia il comportamento operativo atteso del sistema IoT reale, che escluderebbe automaticamente le misurazioni fuori soglia o non conformi alle specifiche dei sensori.

L'analisi comparativa dei modelli di regressione sviluppati e ottimizzati sul dataset simulato ha permesso di valutare, in via preliminare, l'efficacia di diverse architetture di intelligenza artificiale per la previsione della *shelf life* della vinaccia. I risultati riportati si riferiscono a dati utilizzati per la validazione metodologica dei modelli nell'ambito della sperimentazione simulata, finalizzata al confronto tra approcci e non alla stima delle prestazioni su dati reali o di test indipendenti. Sebbene i dati utilizzati non derivino da misure reali, i risultati ottenuti forniscono indicazioni significative per l'individuazione del modello più adatto all'integrazione nel sistema IoT di monitoraggio.

La regressione lineare multipla (MLR) ha rappresentato il modello di riferimento, con prestazioni inferiori ($R^2 = 0,775$, MAE = 3,43, MSE = 17,9), ma utile come *baseline* per quantificare il miglioramento ottenuto con modelli più complessi.

La rete neurale *Multilayer Perceptron Regressor* (MLPRegressor) implementata in Keras, con due strati nascosti, ha mostrato le migliori prestazioni complessive ($R^2 = 0,864$, MAE = 2,33, MSE = 9,22), evidenziando una spiccata capacità di apprendere relazioni non lineari tra le variabili ambientali (temperatura, umidità, pH e luminosità) e la *shelf life*.

La *Support Vector Regression* (SVR) si colloca al secondo posto ($R^2 = 0,858$, MAE = 2,27, MSE = 11,3), mostrando un buon equilibrio tra accuratezza e stabilità. Tale comportamento la rende idonea a contesti con dataset di dimensioni moderate e *features* correttamente scalate.

Il MLPRegressor implementato con scikit-learn, ha raggiunto risultati simili ($R^2 = 0,851$, MAE = 2,48, MSE = 11,8), confermando l'efficacia delle reti neurali *multilayer* anche in scenari tabellari, sebbene con prestazioni leggermente inferiori rispetto alla rete Keras, probabilmente a causa di una minore ottimizzazione architetturale.

I modelli ad albero, *Gradient Boosting Regressor* (GBR) ($R^2 = 0,811$, MAE = 2,68, MSE = 15,1) e *Random Forest Regressor* (RFR) ($R^2 = 0,798$, MAE = 2,57 e MSE = 16,1), hanno fornito risultati coerenti ma non superiori alle reti neurali, suggerendo che le relazioni tra le *features* ambientali e la *shelf life* non siano puramente additive, ma più complesse e continue.

La *Improved ResNet*, architettura convoluzionale progettata *ad hoc*, ha ottenuto risultati inferiori rispetto ai modelli più semplici ($R^2 = 0,786$, MAE = 3,17, MSE = 17,0). Ciò è attribuibile principalmente alla dimensione ridotta del dataset simulato e alla natura non sequenziale dei dati, che limita l'efficacia delle convoluzioni 1D.

In sintesi, l'analisi ha permesso di individuare nella rete neurale Keras il candidato più promettente per una futura integrazione nel sistema IoT, grazie alla sua capacità di modellare relazioni non lineari e di garantire una previsione accurata della *shelf life* in condizioni ambientali variabili.

Tabella III. Risultati dell'analisi comparativa delle prestazioni dei modelli di ML sul *dataset A* (senza antocianine), ottenuti sui dati di validazione della sperimentazione simulata.

Modello	R^2	MAE	MSE
MLPRegressor in Keras	0,879	2,29	9,62
SVR	0,858	2,27	11,3
MLPRegressor in scikit-learn	0,851	2,49	11,8
GBR	0,811	2,68	15,1
Improved ResNet	0,786	3,17	17,0
RFR	0,798	2,57	16,1
MLR	0,775	3,43	17,9

Confronto metriche dei modelli di Intelligenza Artificiale selezionati (dataset A)

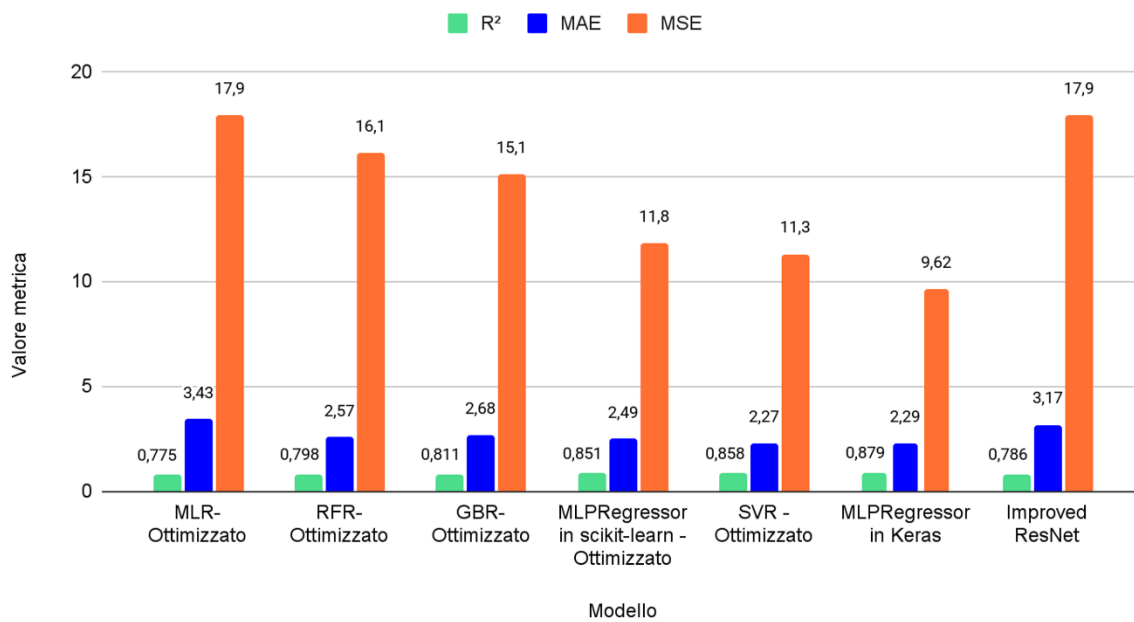


Figura 4.2 Visualizzazione grafica del confronto delle metriche R², MAE, MSE dei modelli sul *dataset A* (senza antocianine).

4.3 ANALISI ESPLORATIVA SULLA VARIABILE “ANTOCIANINE”

All'interno di questa indagine preliminare, finalizzata all'integrazione del modulo di intelligenza artificiale nel dispositivo IoT progettato dall'azienda, si è deciso di valutare l'inclusione di un ulteriore parametro: il livello di antocianine, considerato un indicatore diretto della qualità della vinaccia.

La scelta di introdurre i livelli di antocianine come variabile aggiuntiva deriva dal fatto che questi composti rappresentano uno dei principali target dell'estrazione, essendo responsabili della bioattività dell'uva e del vino, nonché della colorazione caratteristica della biomassa e del vino rosso. Monitorarne l'andamento nel tempo consente di valutare la loro degradazione e, di conseguenza, la stabilità chimica della vinaccia, fornendo un'informazione chiave per la determinazione della *shelf life* e per la valutazione complessiva della qualità del prodotto.

Per questo motivo è stato creato un nuovo dataset di 785 campioni e 7 variabili, chiamato *dataset B*, sviluppato mantenendo le stesse condizioni sperimentali del dataset originario ma includendo anche i livelli di antocianine come parametro aggiuntivo. Tale dataset è stato oggetto di un'analisi esplorativa preliminare per comprendere l'andamento della nuova variabile e la sua relazione con i principali fattori ambientali di conservazione (temperatura,

umidità, pH e luce), al fine di valutarne la rilevanza come indicatore della qualità e della stabilità della vinaccia nel tempo.

4.3.1 STATISTICHE DESCRITTIVE E DISTRIBUZIONE

Dall'analisi delle distribuzioni e dei valori medi delle variabili emerge che le condizioni sperimentali impostate (temperatura, umidità relativa e pH) sono risultate stabili e adeguate per mantenere un buon livello di conservazione del campione (**Figura 4.3**).

In media, le condizioni di conservazione si collocano attorno ai 25,0 °C con una deviazione standard di circa 4 °C, mentre l'umidità relativa presenta una media del 55,8% con una variabilità moderata ($\sigma \approx 11,9\%$). L'intensità luminosa mostra invece una distribuzione molto ampia, compresa tra 0 e 20 000 lux (media $\approx 7\ 200$ lux), ipotizzando che i campioni sono stati esposti a condizioni di luce eterogenee. Il pH varia tra 6,0 e 7,0, con una media di 6,50, corrispondente a un ambiente leggermente acido e chimicamente stabile.

Il livello medio di antocianine è pari a 76,7 a.u. ($\sigma = 5,47$), con una distribuzione quasi normale, segno che la maggior parte dei campioni ha mantenuto un buon contenuto fenolico durante il periodo di osservazione. La *shelf life* media è di circa 103 giorni, con valori compresi tra 88 e 120 giorni, e mostra una distribuzione asimmetrica positiva (o "a coda sinistra") caratterizzata da un forte accumulo di campioni al limite massimo di 120 giorni. Ciò suggerisce che molti campioni erano ancora stabili al termine del test e che la loro durata reale di conservazione potrebbe essere superiore al periodo sperimentale osservato.

La maggior parte dei campioni è rimasta in condizioni di luce contenuta, mentre un gruppo consistente è stato esposto a luce massima (20 000 lux), per testare l'effetto della luce estrema sulla stabilità delle antocianine.

Nel complesso, i risultati indicano che le condizioni di pH, temperatura e umidità scelte sono idonee a garantire una lunga *shelf life*, limitando la degradazione dei pigmenti e mantenendo la qualità complessiva del prodotto. Le antocianine risultano quindi conservate in condizioni ottimali di stabilità chimica, coerenti con un ambiente leggermente acido e termicamente controllato.

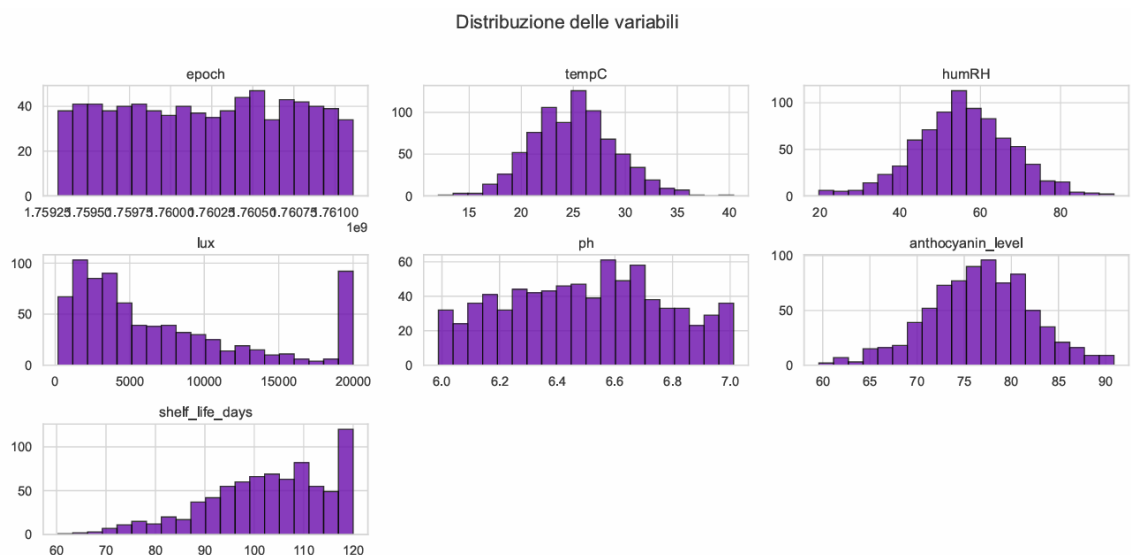


Figura 4.3 Statistiche descrittive e distribuzione delle variabili tempo, temperatura, umidità relativa, luce, pH, livello di antocianine e *shelf life*.

4.3.2 INTERPRETAZIONE DELLE CORRELAZIONI

L'analisi delle correlazioni non si limita a valutare la relazione tra antocianine e *shelf life*, poiché non era noto a priori quali fattori ambientali potessero influenzare la stabilità dei pigmenti nelle condizioni testate. Correlare i livelli di antocianine con temperatura, umidità relativa, luce e pH consente di identificare i fattori che accelerano o rallentano la degradazione dei pigmenti e di ottimizzare le condizioni sperimentali per massimizzare la *shelf life* e la qualità del campione. Allo stesso modo, correlare la *shelf life* con tutte le variabili permette di evidenziare quali fattori influenzano direttamente la durata di conservazione, indipendentemente dai livelli di antocianine.

È importante sottolineare che i risultati ottenuti dalle correlazioni valgono solo per le condizioni sperimentali effettivamente testate (pH, temperatura, umidità e luce) e confermate come idonee alla conservazione delle antocianine; non è possibile generalizzare queste relazioni al di fuori del *range* sperimentale.

L'analisi di correlazione tra le variabili, riportata in **Figura 4.4**, mostra che il livello di antocianine è fortemente correlato con la *shelf life* ($r = 0,85$), indicando che campioni con un contenuto maggiore di antocianine tendono a conservare la qualità più a lungo. Le antocianine risultano influenzate negativamente da temperatura ($r = -0,43$), umidità relativa ($r = -0,53$) e intensità luminosa ($r = -0,44$), mentre pH ($r = -0,14$) e tempo di misurazione ($r = -0,01$) hanno un effetto trascurabile.

Analogamente, la *shelf life* mostra una forte correlazione positiva con le antocianine e correlazioni negative con luce ($r = -0,62$), umidità ($r = -0,54$) e temperatura ($r = -0,39$). pH ($r = -0,07$) e tempo ($r = -0,04$) non sembrano influenzare significativamente la durata di conservazione dei campioni.

Questi risultati indicano che, per garantire una lunga *shelf life* e preservare il contenuto di antocianine, è fondamentale mantenere condizioni di temperatura moderata, umidità controllata e luce ridotta, mentre il pH rimane un fattore poco determinante all'interno dell'intervallo sperimentale osservato. La forte correlazione tra antocianine e *shelf life* conferma inoltre il ruolo cruciale dei pigmenti fenolici nella stabilità chimica e nella conservazione del prodotto.

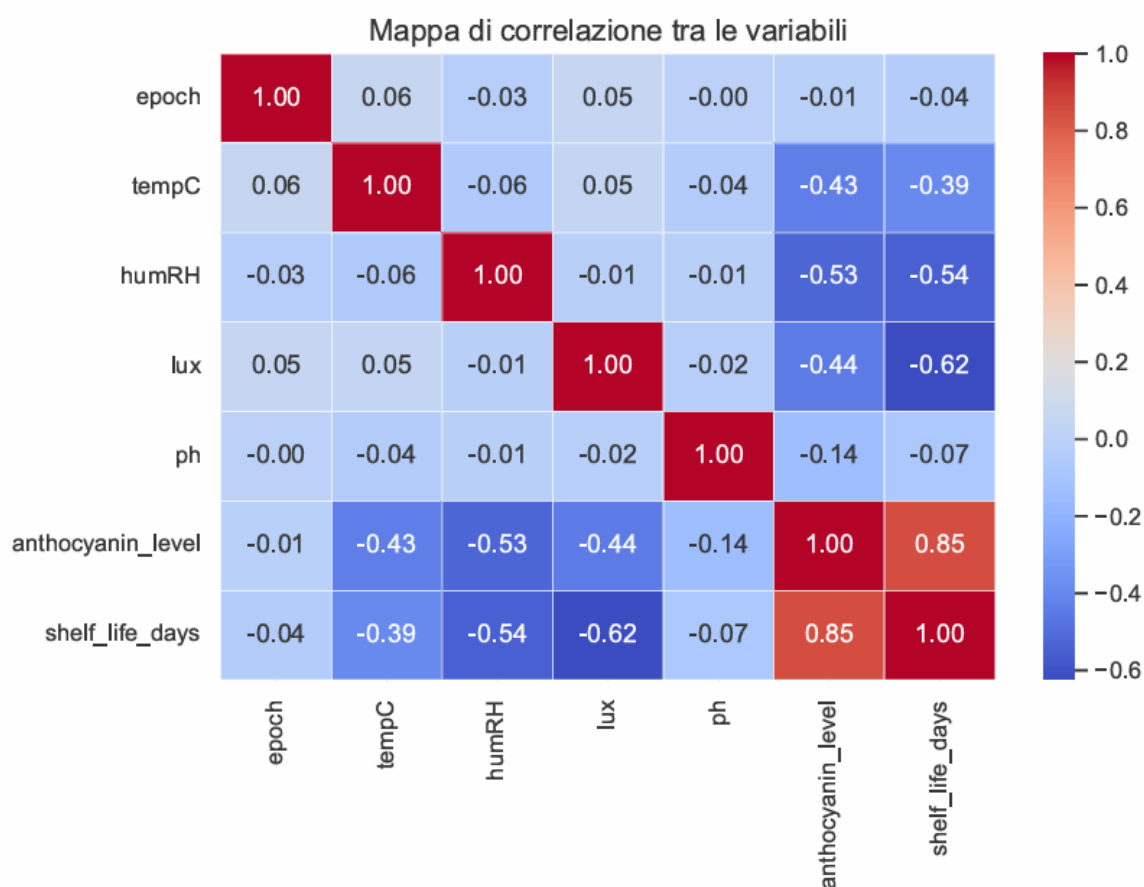


Figura 4.4 Mappa di correlazione fra le variabili selezionate.

Invece, lo scatterplot che mette in relazione la *shelf life* con i livelli di antocianine in **Figura 4.5** mostra chiaramente una tendenza positiva: campioni con contenuto maggiore di antocianine tendono a conservare la qualità più a lungo, confermando la forte correlazione osservata ($r = 0,85$). La colorazione dei punti in funzione della temperatura evidenzia che le temperature moderate (circa 15–30 °C) favoriscono sia il mantenimento di un alto contenuto di antocianine sia una *shelf life* più lunga, con diversi campioni che raggiungono il massimo della durata

osservata (120 giorni). Al contrario, temperature più elevate o estreme risultano associate a una riduzione della *shelf life*, suggerendo un'accelerazione della degradazione dei pigmenti e della qualità complessiva del campione.

Questi risultati confermano il ruolo cruciale delle antocianine nella stabilità chimica e nella conservazione del prodotto e sottolineano l'importanza di controllare la temperatura come leva operativa principale per massimizzare la durata di conservazione. Lo scatterplot, quindi, non solo visualizza la relazione tra pigmenti e *shelf life*, ma integra anche l'effetto dei fattori ambientali, fornendo indicazioni pratiche per l'ottimizzazione delle condizioni di conservazione.

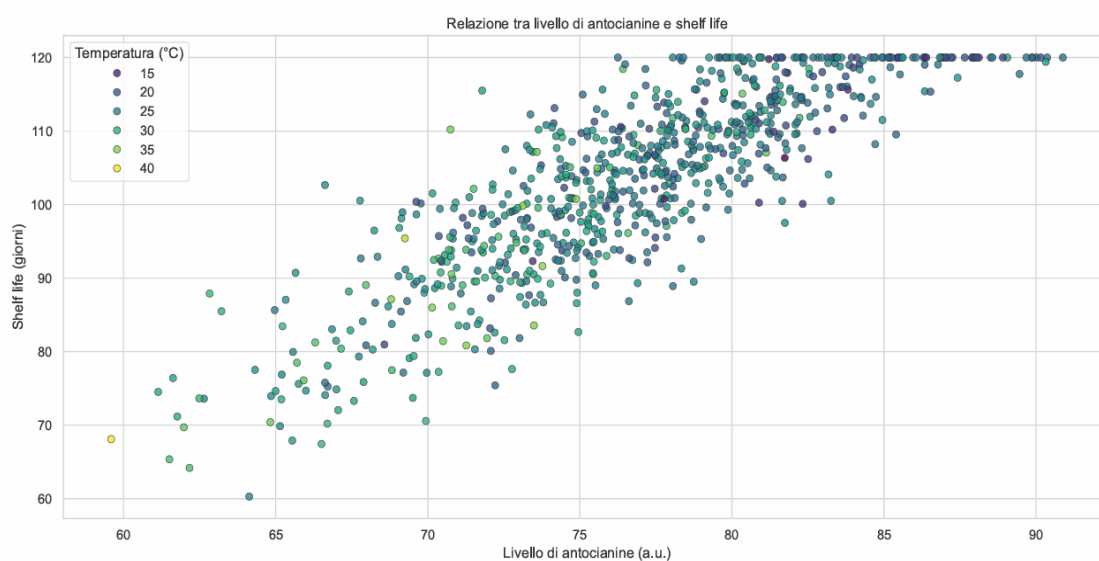


Figura 4.5 Scatterplot della relazione fra il livello di antocianine e *shelf life*.

Infine, per completare l'analisi esplorativa e comprendere meglio le relazioni tra le variabili principali, è stato realizzato un pairplot, che permette di visualizzare simultaneamente l'andamento di ciascuna variabile rispetto alle altre (**Figura 4.6**). Dall'analisi emerge chiaramente la correlazione positiva tra i livelli di antocianine e la *shelf life*, confermando che campioni con un contenuto fenolico maggiore tendono a mantenere la qualità più a lungo. Al contrario, antocianine e temperatura mostrano una relazione inversa lineare: i livelli di pigmento sono più elevati a temperature moderate e diminuiscono con l'aumentare della temperatura. Un andamento simile si osserva per l'umidità relativa, con livelli maggiori di umidità associati a una riduzione degli antociani.

Per pH e intensità luminosa non emergono *pattern* evidenti: i dati risultano più dispersi, suggerendo che entro l'intervallo sperimentale queste variabili hanno un ruolo meno determinante. Le distribuzioni marginali lungo la diagonale confermano le osservazioni

precedenti sugli istogrammi: antocianine, *shelf life*, temperatura e umidità presentano picchi e asimmetrie, mentre pH e luce mostrano maggiore eterogeneità.

In sintesi, il pairplot rafforza le conclusioni delle analisi precedenti, evidenziando le correlazioni chiave tra antocianine, *shelf life* e fattori ambientali, e fornendo indicazioni utili per definire le condizioni ottimali di conservazione dei campioni.

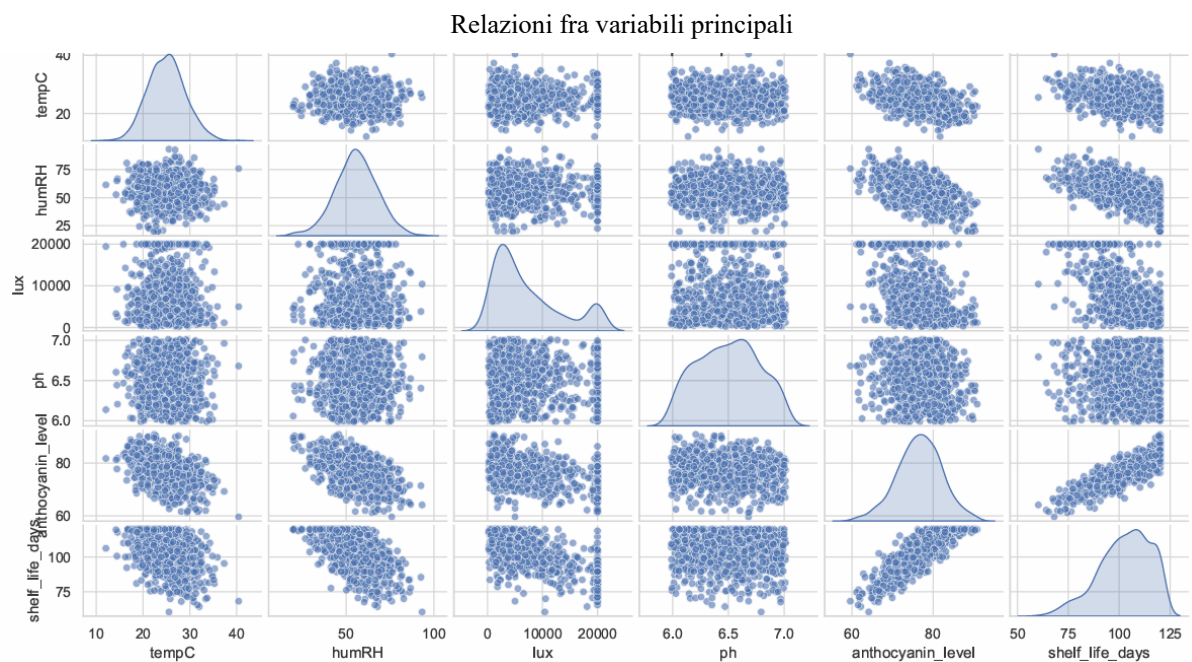


Figura 4.6 Pairplot dell'andamento di ciascuna variabile rispetto alle altre.

4.4 INTEGRAZIONE DEL MODULO ANTOCIANINE E NUOVI RISULTATI (DATASET B)

Per valutare l'effetto dell'inclusione dei livelli di antocianine come nuova variabile predittiva, i modelli di *machine learning* precedentemente addestrati sono stati nuovamente testati sul nuovo dataset di 785 campioni e 7 variabili, il *dataset B*, nell'ambito della stessa procedura di validazione comparativa su dati simulati. L'aggiunta di questa variabile ha comportato un miglioramento significativo delle prestazioni dei modelli di *machine learning* precedentemente testati, evidenziando la rilevanza delle antocianine come indicatore della stabilità e della *shelf life* della vinaccia.

In particolare, la MLPRegressor implementato in scikit-learn e ottimizzato ha raggiunto $R^2 = 0,881$, $MAE = 1,44$ e $MSE = 4,14$, mostrando un buon equilibrio tra accuratezza e stabilità. La SVR ottimizzata ha ottenuto il $MAE = 0,970$ con $R^2 = 0,848$, mentre la rete neurale MLPRegressor in Keras semplice con due strati nascosti ha registrato $R^2 = 0,890$, $MAE = 1,23$

e $MSE = 3,81$, confermandosi leggermente superiore nella capacità di spiegare la varianza complessiva della *shelf life*.

Anche gli altri modelli hanno mostrato un netto miglioramento rispetto al dataset senza antocianine: RFR base ha raggiunto $R^2 = 0,844$, $MAE = 1,12$ e $MSE = 5,42$, migliorando ulteriormente con l'ottimizzazione a $R^2 = 0,851$, $MAE = 1,10$ e $MSE = 5,18$; il GBR base ha ottenuto $R^2 = 0,818$, $MAE = 1,25$ e $MSE = 6,33$, mentre la versione ottimizzata ha mostrato $R^2 = 0,799$, $MAE = 1,19$ e $MSE = 6,98$; infine, la MLR ha registrato $R^2 = 0,585$, $MAE = 2,98$ e $MSE = 14,4$, confermando il suo ruolo di *baseline*. L'architettura convoluzionale *Improved ResNet*, invece, non è migliorata con l'inclusione delle antocianine e ha mostrato prestazioni inferiori ($R^2 = 0,306$, $MAE = 4,19$ e $MSE = 24,1$), probabilmente a causa della natura non sequenziale dei dati e della dimensione limitata del dataset.

In sintesi, l'inclusione dei livelli di antocianine come *feature* ha migliorato le prestazioni complessive di tutti i modelli testati. Tra i modelli più promettenti, la MLPRegressor in Keras semplice rimane la scelta migliore se si vuole massimizzare la spiegazione della varianza, mentre la SVR ottimizzata risulta più adatta se l'obiettivo è minimizzare l'errore medio per campione.

Confronto metriche dei modelli di Intelligenza Artificiale selezionati (dataset B)

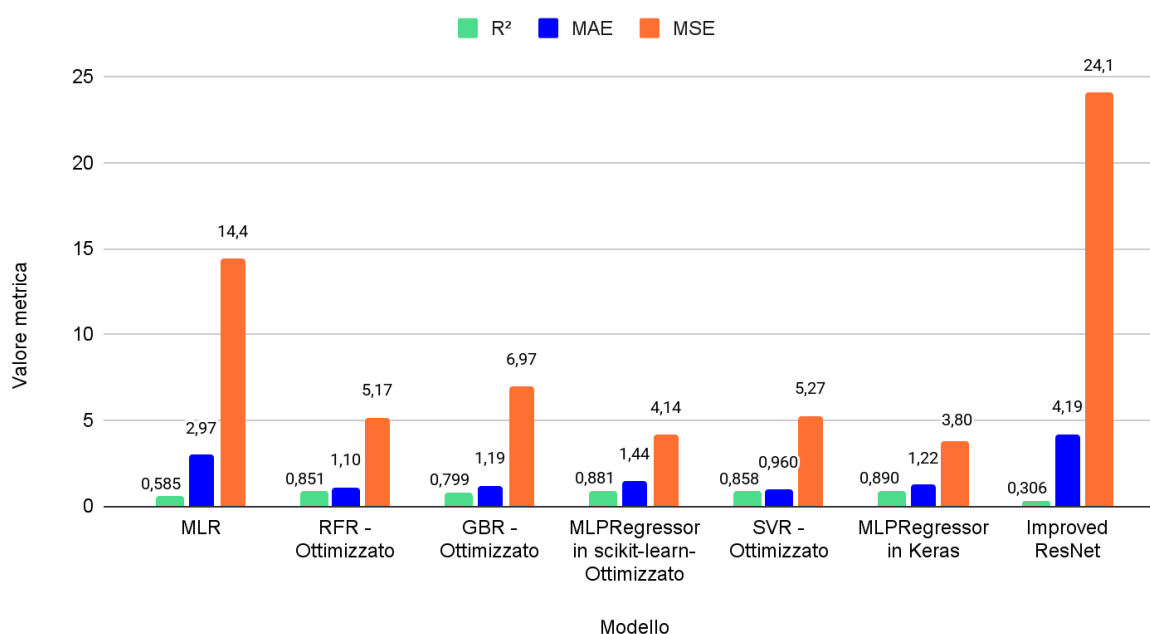


Figura 4.7 Visualizzazione grafica del confronto delle metriche R^2 , MAE, MSE dei modelli sul *dataset B* (con antocianine), a seguito dell'ottimizzazione dei modelli.

Tabella IV. Risultati dell'analisi comparativa delle prestazioni dei modelli di ML sui dataset senza (-) e con (+) antocianine, ottenuti sui dati di validazione della sperimentazione simulata.

Modello	R ²		MAE		MSE	
	-	+	-	+	-	+
MLR	0,775	0,585	3,43	2,97	17,9	14,4
RFR - Ottimizzato	0,798	0,851	2,57	1,10	16,1	5,17
GBR - Ottimizzato	0,811	0,799	2,68	1,19	15,1	6,97
MLPRegressor in scikit-learn-Ottimizzato	0,851	0,881	2,49	1,44	11,9	4,14
SVR - Ottimizzato	0,858	0,858	2,27	0,96	11,3	5,27
MLPRegressor in Keras	0,860	0,890	2,33	1,22	9,22	3,80
Improved ResNet	0,780	0,306	3,17	4,19	17,0	24,1

4.5 ANALISI BOOTSTRAP COMPARATIVA

A seguito del confronto tra i diversi modelli di regressione, si è deciso di approfondire in modo inferenziale l'effetto dell'inclusione della variabile relativa alle antocianine mediante un'analisi bootstrap comparativa.

Questo approccio è stato adottato per quantificare la significatività statistica delle differenze osservate tra i modelli con e senza tale variabile, andando oltre il semplice confronto descrittivo delle metriche di *performance*.

Sono stati pertanto eseguiti numerosi riaddestramenti casuali dei modelli, calcolando a ogni iterazione la differenza tra le metriche di performance (R², MAE e MSE) ottenute sui due dataset. Il *p-value bootstrap* è stato stimato come la proporzione di iterazioni in cui la differenza risultava minore o uguale a zero, fornendo così una misura diretta della probabilità che il miglioramento osservato fosse dovuto al caso.

Le prime analisi sono state condotte sul dataset originale, filtrato in un ristretto intervallo di pH. Tuttavia, la limitata variabilità e la scarsa numerosità campionaria hanno reso tale base dati poco adatta a una valutazione statistica robusta: i campioni generati per ri-campionamento risultavano infatti fortemente sovrapposti e scarsamente rappresentativi della reale variabilità sperimentale.

Per ovviare a questa limitazione, è stata introdotta una *data augmentation* numerica, simulando la variabilità naturale delle misure di laboratorio attraverso l'aggiunta di rumore uniforme casuale sulle variabili di input (temperatura, umidità relativa, illuminamento, pH e livello di antocianine).

In pratica, ogni valore numerico è stato moltiplicato per un fattore casuale compreso tra 0.85–1.15 (rumore $\pm 15\%$) o 0.75–1.25 (rumore $\pm 25\%$), generando nuove configurazioni realistiche dei dati.

Questa tecnica, comunemente adottata nel *machine learning* per incrementare la diversità dei campioni in dataset di dimensioni ridotte, ha consentito di mantenere la struttura del segnale originale pur introducendo la variabilità necessaria a un test statistico affidabile. In questa fase non è stato applicato alcun filtro sul pH, al fine di massimizzare la rappresentatività del campione.

Le Tabelle V, VI, VII riportano le differenze medie di R^2 (ΔR^2), MAE e MSE tra i modelli con e senza antocianine, con i relativi p-value.

Tabella V. Valutazione bootstrap (25% rumore, 1000 iterazioni) delle prestazioni dei modelli: ΔR^2 , ΔMAE , ΔMSE e relativi p-value. Sono evidenziati in grassetto i p-value significativi ed è presente un commento esplicativo dei risultati.

Modello	ΔR^2	p-value (ΔR^2)	ΔMAE	p-value (ΔMAE)	ΔMSE	p-value (ΔMSE)	Sintesi
MLR	+0.110	0.105	-0.68	0.072	-6.99	0.196	Miglioramento coerente ma non significativo
RFR	+0.050	0.235	-0.74	0.066	-5.72	0.258	Miglioramento lieve, p vicino alla soglia
GBR	+0.060	0.227	-0.820	0.040	-7.02	0.184	Miglioramento significativo su MAE
MLPRegressor in scikit-learn	+0.210	0.030	-0.190	0.331	+2.97	0.464	R^2 significativamente più alto, ma errori incoerenti
SVR	+0.100	0.141	-1.13	0.031	-8.21	0.195	Miglioramento consistente, significativo per MAE
MLPRegressor in Keras	+0.122	0.084	-0.616	0.119	-7.318	0.208	Miglioramento <i>borderline</i> ($p \approx 0.08$)

Tabella VI. Valutazione bootstrap (15% rumore, 1000 iterazioni) delle prestazioni dei modelli: ΔR^2 , ΔMAE , ΔMSE e relativi p-value. Sono evidenziati i p-value significativi ed è presente un commento esplicativo dei risultati.

Modello	ΔR^2	p-value (ΔR^2)	ΔMAE	p-value (ΔMAE)	ΔMSE	p-value (ΔMSE)	Sintesi
MLR	+0.090	0.064	-0.346	0.194	-4.16	0.226	Miglioramento moderato ma non significativo
RFR	+0.029	0.320	-0.802	0.030	-5.26	0.214	MAE significativamente migliore; altri effetti non significativi
GBR	+0.029	0.320	-0.832	0.012	-6.31	0.144	Miglioramento significativo su MAE; benefici coerenti sulle altre metriche
SVR	+0.068	0.148	-0.948	0.028	-5.28	0.238	Riduzione significativa del MAE; miglioramenti non significativi su R^2 e MSE
MLPRegressor in scikit-learn	+0.217	0.030	+0.238	0.470	+7.289	0.390	R^2 significativamente più alto, ma peggioramento degli errori
MLPRegressor in Keras	+0.105	0.062	-0.248	0.327	-1.809	0.385	Miglioramento coerente ma p-value non significativi

Tabella VII. Risultati del bootstrap per la rete Fully Connected: differenze di R^2 (ΔR^2), MAE (ΔMAE), e MSE (ΔMSE) con relativi p-value (rumore 15% e 25%).

Rumore \pm	Iterazioni	ΔR^2	p-value (ΔR^2)	ΔMAE	p-value (ΔMAE)	ΔMSE	p-value (ΔMSE)
25%	300	+0.138	0.060	-0.519	0.176	-3.83	0.343
	400	+0.108	0.100	-0.621	0.112	-7.51	0.217
	600	+0.125	0.062	-0.425	0.200	-5.28	0.272
	1000	+0.122	0.084	-0.616	0.119	-7.32	0.208
15%	300	+0.082	0.093	-0.425	0.200	-4.16	0.270
	400	+0.077	0.120	-0.490	0.147	-5.92	0.177
	1000	+0.106	0.062	-0.248	0.327	-1.81	0.385

In generale, l'aggiunta della variabile antocianine mostra un miglioramento coerente delle prestazioni predittive, con effetti più evidenti nei modelli non lineari. L'analisi bootstrap ha mostrato una tendenza coerente al miglioramento delle prestazioni predittive in presenza della variabile "antocianine", sebbene con livelli di significatività variabili a seconda del modello e dell'intensità del rumore.

Applicando il rumore del $\pm 15\%$, i risultati evidenziano incrementi medi di ΔR^2 compresi tra +0.03 e +0.22, con riduzioni di MAE e MSE nella maggior parte dei casi. Le differenze più significative emergono per i modelli RFR (MAE, $p = 0.03$), GBR (MAE, $p = 0.012$), SVR (MAE, $p = 0.028$) e MLPRegressor in scikit-learn (R^2 , $p = 0.03$). Anche i modelli lineari mostrano un miglioramento coerente ($\Delta R^2 = +0.09$; $p = 0.064$), pur non raggiungendo la soglia di significatività statistica.

Con rumore $\pm 25\%$, le differenze si mantengono nella stessa direzione ma con p-value mediamente più alti, riflettendo la maggiore varianza introdotta.

La MLR mostra un incremento limitato di ΔR^2 (+0.11) non significativo ($p = 0.105$), mentre la RFR evidenzia un miglioramento ancora più contenuto ($\Delta R^2 = +0.05$; $p = 0.235$). Il GBR

presenta un aumento moderato di ΔR^2 (+0.06; $p = 0.227$), accompagnato tuttavia da una riduzione significativa del MAE ($\Delta MAE = -0.82$; $p = 0.040$). Un comportamento simile si osserva per la SVR, con una riduzione significativa del MAE ($\Delta MAE = -1.13$; $p = 0.031$) ma un incremento di ΔR^2 non significativo (+0.10; $p = 0.141$). Tra i modelli neurali, l'MLPRegressor in scikit-learn mantiene un miglioramento significativo in termini di ΔR^2 (+0.21; $p = 0.030$).

Per il modello Keras (MLPRegressor), con 1000 iterazioni di bootstrap e rumore $\pm 15\%$, si è osservato un incremento medio di $\Delta R^2 = +0.106$ con $p = 0.062$, accompagnato da una riduzione coerente di MAE (-0.25) e MSE (-1.81), sebbene non statisticamente significative ($p > 0.3$). Con rumore $\pm 25\%$, il modello Keras mantiene la stessa direzione dell'effetto, con un incremento medio di $\Delta R^2 = +0.122$ e un p-value di 0.084, insieme a una riduzione di MAE (-0.62) e MSE (-7.32). Sebbene la significatività statistica risulti leggermente ridotta rispetto al caso $\pm 15\%$, riflettendo la maggiore varianza introdotta dal rumore, l'effetto rimane coerente e positivo, a dimostrazione della robustezza del modello alle fluttuazioni dei dati e con un numero elevato di repliche.

Le distribuzioni bootstrap di ΔR^2 (**Figura 4.8**) risultano spostate verso valori positivi, indicando che nella quasi totalità delle iterazioni il modello con la variabile *anthocyanin_level* spiega una quota di varianza superiore rispetto a quello senza.

Al contrario, le distribuzioni di ΔMAE e ΔMSE appaiono più simmetriche, suggerendo che le differenze negli errori, pur coerenti con un miglioramento delle prestazioni, non sono statisticamente robuste ($p > 0.1$).

In particolare, al crescere del livello di rumore, le distribuzioni mostrano una maggiore dispersione, ma la tendenza al miglioramento resta invariata, con ΔR^2 mediamente positivo e riduzioni di errore coerenti con quanto osservato nel caso $\pm 15\%$.

Nel complesso, i risultati per il modello Keras evidenziano una tendenza stabile e riproducibile al miglioramento delle prestazioni predittive con l'inclusione della variabile "antocianine". L'effetto, pur moderato, risulta consistente su diversi livelli di rumore e in linea con quanto osservato nei modelli non lineari come MLP e SVR, confermando il contributo informativo delle antocianine nella descrizione dei processi di degradazione e nella stima della *shelf life*.

Distribuzioni bootstrap delle differenze (con – senza antocianine)

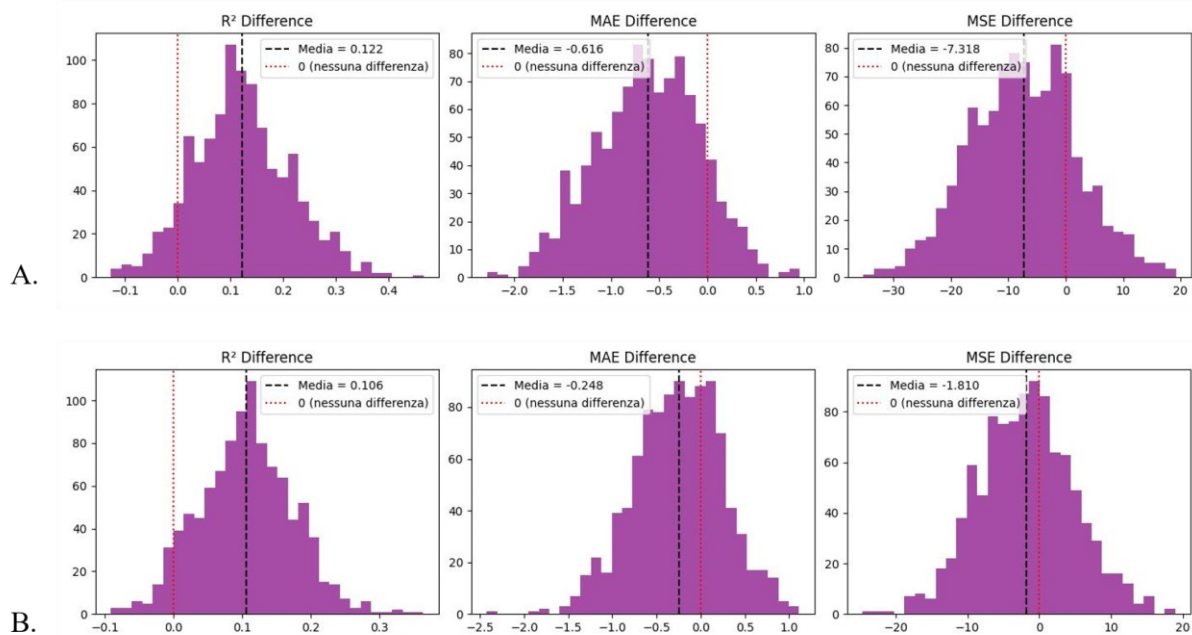


Figura 4.8. Distribuzione Keras, 1000 iterazioni, rumore del 25% (A) e rumore del 15% (B).

In sintesi, la variabile delle antocianine ha mostrato un contributo positivo, coerente e replicabile all'interno di tutte le configurazioni testate, con un effetto più marcato e stabile nei modelli non lineari (MLP, SVR, Keras) rispetto a quelli lineari o ad albero.

Tra tutti i modelli testati, Keras (rete neurale *feed-forward*) ha evidenziato la maggiore stabilità e coerenza nei risultati, sia al variare del rumore sperimentale sia del numero di iterazioni.

In tutti gli esperimenti, l'aggiunta delle antocianine ha portato a un aumento del coefficiente di determinazione R^2 compreso tra +0.07 e +0.14, segnalando una migliore capacità del modello di spiegare la variabilità del dato.

Gli errori medi (MAE e MSE) si sono ridotti in modo coerente, anche se senza raggiungere la significatività statistica convenzionale ($p \approx 0.06-0.12$).

Ciò suggerisce che il miglioramento non sia dovuto al caso, ma che la numerosità e l'eterogeneità dei dati disponibili non siano ancora sufficienti per renderlo pienamente significativo ($p < 0.05$).

Dal punto di vista scientifico, l'effetto delle antocianine appare moderato ma costante: la loro inclusione fornisce al modello informazioni aggiuntive utili a descrivere meglio il processo di degradazione e la durata di conservazione del prodotto.

La consistenza del segnale attraverso livelli di rumore differenti ($\pm 15\%$ e $\pm 25\%$) e l'alta ripetibilità delle differenze positive di R^2 indicano un fenomeno stabile e generalizzabile.

Nel complesso, i risultati bootstrap indicano che l'aggiunta delle antocianine come variabile esplicativa migliora sistematicamente la capacità predittiva dei modelli di regressione, anche in condizioni di maggiore variabilità sperimentale.

Sebbene la significatività statistica risulti *borderline* per diversi modelli (p compreso tra 0.06 e 0.12), l'uniformità della direzione dell'effetto e la coerenza tra modelli differenti costituiscono un'evidenza scientificamente solida di un impatto reale, seppur moderato, delle antocianine sulla qualità predittiva.

Con un dataset più ampio e più eterogeneo, ad esempio con un numero maggiore di campioni sperimentali e una variabilità più rappresentativa delle condizioni reali, si prevede che l'effetto delle antocianine possa emergere in modo statisticamente significativo. Ciò rafforzerebbe ulteriormente la loro interpretazione come variabile chiave nella modellazione della *shelf life* e nella descrizione dei processi di degradazione del prodotto.

5. CONCLUSIONI

Il *grape pomace*, sottoprodotto che si accumula principalmente durante la vinificazione, rappresenta una risorsa biologicamente preziosa e ricca di nutrienti di grande interesse applicativo. Numerosi studi hanno evidenziato come la sua valorizzazione, tramite tecniche avanzate di estrazione e analisi, possa fornire informazioni significative sulla qualità, sulla composizione fenolica e sul potenziale di conservazione delle uve e dei prodotti derivati. In questo contesto, l'integrazione tra sensori innovativi, analisi chimiche e modelli predittivi basati sull'intelligenza artificiale costituisce un'opportunità concreta per migliorare i processi di monitoraggio e sfruttare in modo più efficiente il patrimonio informativo contenuto nei sottoprodotti della filiera vitivinicola.

Il lavoro svolto durante il tirocinio si colloca all'interno di questo scenario di innovazione, con l'obiettivo di valutare la possibilità di potenziare le capacità predittive di un dispositivo aziendale per il monitoraggio a basso impatto sul campione della qualità della vinaccia, basato sulla dispersione controllata della vinaccia in acqua distillata per la misurazione del pH e di altri parametri attraverso sensori dedicati. Tale obiettivo è stato perseguito attraverso l'integrazione di metodologie di analisi dei dati e modelli avanzati di apprendimento automatico, oltre che mediante una revisione critica della letteratura scientifica, che ha evidenziato una progressiva transizione verso approcci predittivi basati su *imaging*, spettroscopia VIS–NIR e tecniche *data-driven*, in linea con gli sviluppi dell'agricoltura 4.0 e con le più recenti prospettive dell'agricoltura 5.0. Questo quadro teorico ha confermato la centralità delle variabili già monitorate dall'azienda e, allo stesso tempo, ha suggerito la possibilità di esplorare parametri addizionali in grado di incrementare la precisione delle stime. In particolare, la maggiore capacità predittiva offerta dall'intelligenza artificiale si rivela uno strumento ideale per identificare i lotti di scarti vitivinicoli che mantengono il miglior profilo polifenolico nel tempo. Ciò consente di ottimizzare l'efficienza e la qualità dei processi di valorizzazione sostenibile, favorendo l'adozione di strategie innovative di riutilizzo in settori quali alimentazione, cosmetica, farmaceutica e agronomia.

Sulla base di queste considerazioni, è stata condotta un'analisi preliminare dei principali algoritmi di regressione, comprendendo modelli lineari, metodi ensemble, *Support Vector Regression* e architetture di *deep learning*, sia classiche (MLP) sia avanzate (reti *fully connected* implementate in Keras). I modelli sono stati valutati sia sui dati originari sia su una versione estesa del dataset contenente la variabile "antocianine", con l'obiettivo di quantificare il contributo informativo di quest'ultima. Tra i modelli analizzati, la rete neurale *Multilayer Perceptron Regressor* (MLP) implementata in Keras, con due strati nascosti, si è rivelata la

soluzione più promettente, evidenziando le migliori prestazioni complessive. Il modello ha dimostrato, rispetto agli altri approcci considerati, una maggiore efficacia nell'apprendimento di relazioni non lineari tra le variabili ambientali (temperatura, umidità, pH e luminosità) e la *shelf life*, mantenendo risultati stabili e coerenti in tutti gli scenari ipotizzati nel presente studio. Le analisi esplorative hanno evidenziato associazioni rilevanti tra il contenuto di antocianine e la *shelf life*, suggerendo la possibile utilità di questa variabile come indicatore predittivo. Tale ipotesi è stata approfondita attraverso un'ampia campagna di sperimentazioni, basata su procedure di bootstrap fino a 1000 iterazioni e sull'introduzione controllata di rumore ($\pm 15\%$ e $\pm 25\%$), al fine di simulare condizioni operative realistiche e testare la robustezza dei modelli. I risultati ottenuti mostrano che l'inclusione della variabile "antocianine" determina un miglioramento sistematico delle prestazioni predittive nella maggior parte dei modelli, con effetti più marcati nelle architetture non lineari (MLP, SVR, GBR). In particolare, le reti neurali sviluppate in Keras hanno mostrato incrementi stabili di ΔR^2 compresi tra +0.10 e +0.14 nelle diverse configurazioni testate, accompagnati da una riduzione coerente degli errori MAE e MSE. Pur non raggiungendo sempre la soglia di significatività statistica convenzionale ($p < 0.05$), la consistenza e la ripetibilità dei risultati, anche in presenza di rumore, suggeriscono che il contributo delle antocianine non sia attribuibile al caso, ma rifletta un effetto reale e potenzialmente sfruttabile nel contesto applicativo.

Nel complesso, il lavoro svolto ha permesso di:

- individuare con maggiore chiarezza le variabili più informative per il dispositivo aziendale;
- identificare modelli predittivi adatti all'integrazione in sistemi *embedded*;
- valutare quantitativamente l'apporto della variabile "antocianine" sulla predizione della *shelf life*;
- definire una base metodologica solida per future ottimizzazioni del dispositivo di monitoraggio.

In prospettiva, il passo successivo prevede l'integrazione dei modelli predittivi più performanti all'interno del dispositivo aziendale, trasformandolo in un sistema di monitoraggio intelligente capace di stimare automaticamente la *shelf life* della vinaccia. Grazie alla struttura modulare del backend, già predisposto per l'invocazione dei modelli predittivi tramite richieste HTTP (*Hypertext Transfer Protocol*) o script Python, l'implementazione di un modulo di intelligenza artificiale rappresenta un'evoluzione concreta e relativamente agevole, senza necessità di modifiche alle componenti hardware o al firmware principale.

Per supportare questa integrazione e migliorarne l'affidabilità, sarà fondamentale ampliare il dataset mediante nuove campagne sperimentali, condotte sia in condizioni controllate sia in

contesti reali, così da aumentare la generalizzabilità dei modelli e consolidare le evidenze già ottenute sul contributo delle variabili informative, come le antocianine, alla predizione della *shelf life*. I risultati preliminari indicano che l'inclusione delle antocianine come variabile aggiuntiva migliora in modo sistematico le *performance* predittive, soprattutto nei modelli non lineari o nelle reti neurali, fornendo un parametro scientificamente rilevante per comprendere i processi di degradazione della vinaccia.

L'integrazione di tecniche avanzate di *feature extraction*, in particolare derivate da sistemi di *imaging*, potrebbe ulteriormente potenziare le capacità predittive del dispositivo. Inoltre, l'implementazione dei modelli più performanti in dispositivi *embedded* a basso consumo energetico costituirebbe un passo decisivo verso un sistema di monitoraggio intelligente pienamente integrabile nei processi produttivi aziendali.

Infine, la flessibilità del dispositivo e dei modelli sviluppati suggerisce che lo stesso approccio possa essere esteso ad altre tipologie di biomassa deperibile, aprendo nuove opportunità per la valorizzazione sostenibile dei sottoprodotti agroalimentari pienamente integrabile nei processi produttivi aziendali e potenzialmente applicabile nell'ambito dell'agricoltura 4.0 e delle più recenti visioni dell'agricoltura 5.0.

6. BIBLIOGRAFIA

1. Abbate, S., Centobelli, P., & Di Gregorio, M. (2025). Wine waste valorisation: Crushing the research domain. *Review of Managerial Science*, 19, 963–998. <https://doi.org/10.1007/s11846-024-00779-5>
2. Abdelbasset, W. K., Elkholi, S. M., Opulencia, M. J. C., Diana, T., Su, C.-H., Alashwal, M., Zwawi, M., Algarni, M., Abdelrahman, A., & Nguyen, H. C. (2022). Development of multiple machine-learning computational techniques for optimization of heterogeneous catalytic biodiesel production from waste vegetable oil. *Arabian Journal of Chemistry*, 15(6), 103843. <https://doi.org/10.1016/j.arabjc.2022.103843>
3. Abreu, T., Sousa, P., Gonçalves, J., Hontman, N., Teixeira, J., Câmara, J. S., & Perestrelo, R. (2024). Grape Pomace as a Renewable Natural Biosource of Value-Added Compounds with Potential Food Industrial Applications. *Beverages*, 10(2), 45. <https://doi.org/10.3390/beverages10020045>
4. Adamashvili, N., Zhizhilashvili, N., & Tricase, C. (2024). The integration of the Internet of Things, artificial intelligence, and blockchain technology for advancing the wine supply chain. *Computers*, 13(3), 72. <https://doi.org/10.3390/computers13030072>
5. Alegbeleye, O., Odeyemi, O. A., Strateva, M., & Stratev, D. (2022). Microbial spoilage of vegetables, fruits and cereals. *Applied Food Research*, 2(1), 100122. <https://doi.org/10.1016/j.afres.2022.100122>
6. Alizadeh Behbahani, B., Jooyandeh, H., Taki, M., & Falah, F. (2024). Evaluation of the probiotic, anti-bacterial, anti-biofilm, and safety properties of *Lactocaseibacillus paracasei* B31-2. *LWT*, 207, 116676. <https://doi.org/10.1016/j.lwt.2024.116676>
7. Allaw, M., Perra, M., Parekh, P., Serra, M., Marongiu, J., Castangia, I., Fulgheri, F., Caboni, P., Tolle, G., Corrias, F., Asunis, F., Giua, G., Mohammad, F., Ghavam, M., Morelli, M., Manconi, M., & Manca, M. L. (2025). Antioxidant and neuroprotective effects of nutriosomes and grape pomace phytochemicals in a cell model of oxidative stress and mouse model of Parkinson disease. *Scientific Reports*, 15, 11947. <https://doi.org/10.1038/s41598-025-95060-4>
8. Allegro, G., Pastore, C., Valentini, G., Muzzi, E., & Filippetti, I. (2016). Influence of berry ripeness on accumulation, composition and extractability of skin and seed

- flavonoids in cv. *Sangiovese* (*Vitis vinifera* L.). *Journal of the Science of Food and Agriculture*, 96(13), 4553–4559. <https://doi.org/10.1002/jsfa.7672>
9. Aning-Dei, E., Yu, J., & Ibrahim, S. A. (2025). Detoxification of Grape Pomace Contaminated with Ochratoxin A by Thermal–Pressure Treatment in Combination with Lactic Acid Bacteria Fermentation. *Microorganisms*, 13(9), 1972. <https://doi.org/10.3390/microorganisms13091972>
 10. Ascrizzi, R., Pieracci, Y., Melai, B., Cioni, P. L., Narri, P., Flamini, G. & Pistelli, L. (2022). Preliminary Study on Red Wine Aroma: The Volatile Profiles of Six Grape Cultivars in Different Vinification Phases. *Fermentation*, 8(12), 753. <https://doi.org/10.3390/fermentation8120753>
 11. Awad, M., & Khanna, R. (2015). Support vector regression. *Efficient learning machines* (pp. 67–80). Apress. https://doi.org/10.1007/978-1-4302-5990-9_4
 12. Avila, F., Mora, M., & Fredes, C. (2014). A method to estimate grape phenolic maturity based on seed images. *Computers and Electronics in Agriculture*, 101, 76–83. <https://doi.org/10.1016/j.compag.2013.12.006>
 13. Baiano, A. (2021). An Overview on Sustainability in the Wine Production Chain. *Beverages*, 7(1), 15. <https://doi.org/10.3390/beverages7010015>
 14. Balaine, L., Läßle, D., Dillon, E. J., & Buckley, C. (2023). Extension and management pathways for enhanced farm sustainability: Evidence from Irish dairy farms. *European Review of Agricultural Economics*, 50, 810–850. <https://doi.org/10.1093/erae/jbac033>
 15. Barth, M., Hankinson, T. R., Zhuang, H., & Breidt, F. (2009). Microbiological spoilage of fruits and vegetables. In W. H. Sperber & M. P. Doyle (Eds.), *Compendium of the microbiological spoilage of foods and beverages* (pp. 135–183). Springer. https://doi.org/10.1007/978-1-4419-0826-1_6
 16. Brito, C., Pereira, S., Martins, S., Monteiro, A., Moutinho-Pereira, J. M., & Dinis, L. (2024, October 8). *Strategies for achieving the sustainable development goals across the wine chain: A review*. *Frontiers in Sustainable Food Systems*, 8. <https://doi.org/10.3389/fsufs.2024.1437872>
 17. Biesalski, H. K., Dragsted, L. O., Elmadafa, I., Grossklaus, R., Müller, M., Schrenk, D., Walter, P., & Weber, P. (2009). Bioactive compounds: Definition and assessment of activity. *Nutrition*, 25(11–12), 1202–1205. <https://doi.org/10.1016/j.nut.2009.04.023>
 18. Blekos A., Chatzis K., Kotaidou M., Chatzis T., Solachidis V., Konstantinidis D. & Dimitropoulos, K. (2023). *A Grape Dataset for Instance Segmentation and*

Maturity Estimation. Agronomy, 13(8), 1995.
<https://doi.org/10.3390/agronomy13081995>

19. Boulton, R. (2001). The copigmentation of anthocyanins and its role in the color of red wine: A critical review. *American Journal of Enology and Viticulture*, 52(2), 67–87. <https://doi.org/10.5344/ajev.2001.52.2.67>
20. Bustamante, M. A., Moral, R., Paredes, C., Pérez-Espinosa, A., Moreno-Caselles, J., & Pérez-Murcia, M. D. (2008). Agrochemical characterisation of the solid by-products and residues from the winery and distillery industry. *Waste Management*, 28(2), 372–380. <https://doi.org/10.1016/j.wasman.2007.01.013>
21. Carter, L. G., D'Orazio, J. A., & Pearson, K. J. (2014). Resveratrol and cancer: focus on in vivo evidence. *Endocrine-related cancer*, 21(3), R209-R225. <https://doi.org/10.1530/ERC-13-0171>
22. Casini, L., Cavicchi, A., Corsi, A., & Santini, C. (2010). Hopelessly devoted to sustainability: Marketing challenges to face in the wine business. In *Proceedings of the 119th EAAE Seminar “Sustainability in the Food Sector: Rethinking the Relationship between the Agro-Food System and the Natural, Social, Economic and Institutional Environments”*. (Capri, Italy, 30 June–2 July 2010). <https://hdl.handle.net/20.500.12606/31020>
23. Casson, A., Ortuani, B., Giovenzana, V., Brancadoro, L., Corsi, S., Gharsallah, O., Guidetti, R., & Facchi, A. (2022) A multidisciplinary approach to assess environmental and economic impact of conventional and innovative vineyards management systems in Northern Italy. *Science of the Total Environment*, 838, 156181. <https://doi.org/10.1016/j.scitotenv.2022.156181>
24. Castellanos-Gallo, L., Ballinas-Casarrubias, L., Espinoza-Hicks, J. C., Hernández-Ochoa, L. R., Muñoz-Castellanos, L. N., Zermeño-Ortega, M. R., Borrego-Loya, A., & Salas, E. (2022). Grape Pomace Valorization by Extraction of Phenolic Polymeric Pigments: A Review. *Processes*, 10(3), 469. <https://doi.org/10.3390/pr10030469>
25. Chaudhry, M. M. A., Amodio, M. L., Babellahi, F., de Chiara, M. L. V., Amigo Rubio, J. M., & Colelli, G. (2018). *Hyperspectral imaging and multivariate accelerated shelf life testing (MASLT) approach for determining shelf life of rocket leaves*. *Journal of Food Engineering*, 238, 122–133. <https://doi.org/10.1016/j.jfoodeng.2018.06.017>
26. Chen, Y., Deng, J., Liu, Z., Chen, J., Wei, W., Xiang, Y., Zhu, X., & Li, C. (2025). *Grape sugar content prediction with multispectral alignment and improved*

- residual network. Scientific Reports, 15*, 36927. <https://doi.org/10.1038/s41598-025-20848-3>
27. Cheng, Y., Dong, H., Wu, Y. Y., & Xiao, K. J. (2021). Preparation of an amidated graphene oxide/sulfonated poly ether ether ketone (AGO/SPEEK) modified atmosphere packaging for the storage of cherry tomatoes. *Foods, 10*(3), 552. <https://doi.org/10.3390/foods10030552>
 28. Chouchouli, V., Kalogeropoulos, N., Konteles, S., Karvela, E., Makris, D., & Karathanos, V. (2013). *Fortification of yoghurts with grape (Vitis vinifera) seed extracts. LWT – Food Science and Technology, 53*(2), 522–529. <https://doi.org/10.1016/j.lwt.2013.03.008>
 29. Choudhary, A. (2024). Internet of things: A comprehensive overview, architectures, applications, simulation tools, challenges and future directions. *Discovery Internet of Things, 4*, 31. <https://doi.org/10.1007/s43926-024-00084-3>
 30. Conde, C., Silva, P., Natacham, F., Dias, A. C. P., Tavares, R. M., Sousa, M. J., Agasse, A., Delrot, S., & Gerós, H. (2007). Biochemical changes throughout grape berry development and fruit and wine quality. *Food, 1*, 1–22. https://www.researchgate.net/publication/228555202_Biochemical_changes_throughout_Grape_Berry_development_and_fruit_and_wine_quality
 31. Costa, J. M., Catarino, S., Escalona, J. M., & Comuzzo, P. (2022). Achieving a more sustainable wine supply chain—environmental and socioeconomic issues of the industry. In *Improving sustainable viticulture and winemaking practices* (pp. 1–24). Elsevier. <https://doi.org/10.1016/B978-0-323-85150-3.00009-8>
 32. Cui, F., Zheng, S., Wang, D., Tan, X., Li, Q., Li, J., & Li, T. (2023). Recent advances in shelf life prediction models for monitoring food quality. *Comprehensive Reviews in Food Science and Food Safety, 22*(2), 1257-1284. <https://doi.org/10.1111/1541-4337.13110>
 33. Deng, H., Zhou, Y., Wang, L., & Zhang, C. (2021). *Ensemble learning for the early prediction of neonatal jaundice with genetic features. BMC Medical Informatics and Decision Making, 21*, 305. <https://doi.org/10.1186/s12911-021-01701-9>
 34. Di Stefano, V., Buzzanca, C., Melilli, M. G., Indelicato, S., Mauro, M., Vazzana, M., Arizza, V., Lucarini, M., Durazzo, A., & Bongiorno, D. (2022). Polyphenol characterization and antioxidant activity of grape seeds and skins from Sicily: A preliminary study. *Sustainability, 14*, 6702. <https://doi.org/10.3390/su14116702>

35. Dixon, P. M. (2006). Bootstrap resampling. In *Encyclopedia of Environmetrics*. <https://doi.org/10.1002/9780470057339.vab028>
36. Donia, E., Mineo, A. M., & Sgroi, F. (2018). A methodological approach for assessing business investments in renewable resources from a circular economy perspective. *Land Use Policy*, 76, 823–827. <https://doi.org/10.1016/j.landusepol.2018.03.017>
37. Dwivedi, S. K., Roy, P., Karda, C., Agrawal, S., & Amin, R. (2021). Blockchain-based Internet of Things and industrial IoT: A comprehensive survey. *Security and Communication Networks*, 2021, Article 7142048. <https://doi.org/10.1155/2021/7142048>
38. Duan, Y., Wang, G. B., Fawole, O. A., Verboven, P., Zhang, X. R., Wu, D., Opara, U. L., Nicolai, B., & Chen, K. (2020). Postharvest precooling of fruit and vegetables: A review. *Trends in Food Science & Technology*, 100, 278–291. <https://doi.org/10.1016/j.tifs.2020.04.027>
39. Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9780429246593>
40. El Mrabet, Z., Sugunraj, N., Ranganathan, P., & Abhyankar, S. (2022). Random Forest Regressor-Based Approach for Detecting Fault Location and Duration in Power Systems. *Sensors*, 22(2), 458. <https://doi.org/10.3390/s22020458>
41. Expósito, I., Cuiñas, I., & Gay-Fernández, J. A. (2013). Efficient traceability solutions in the wine production by RFID and WSN. In *Proceedings of the European Conference on Antennas and Propagation, Gothenburg, Sweden, 8–12 April 2013*. <https://ieeexplore.ieee.org/document/6546968>
42. Ezeudu, O. B., Agunwamba, J. C., Ezeudu, T. S., Ugochukwu, U. C., & Ezeasor, I. C. (2021). Natural leaf-type as food packaging material for traditional food in Nigeria: Sustainability aspects and theoretical circular economy solutions. *Environmental Science and Pollution Research*, 28, 8833–8843. <https://doi.org/10.1007/s11356-020-11268-z>
43. Fanourakis, D., Pieruschka, R., Savvides, A., Macnish, A. J., Sarlikioti, V., & Woltering, E. J. (2013). Sources of vase life variation in cut roses: A review. *Postharvest Biology and Technology*, 78, 1–15. <https://doi.org/10.1016/j.postharvbio.2012.12.001>
44. Ferro, M. V., & Catania, P. (2023). Technologies a Innovative Methods for Precision Viticulture: A Comprehensive Review. *Horticulturae*, 9(3), 399. <https://doi.org/10.3390/horticulturae9030399>

45. Ferrer, J. R., García-Cortijo, M. C., Pinilla, V., & Castillo-Valero, J. S. (2022). The business model and sustainability in the Spanish wine sector. *Journal of Cleaner Production*, 330, 129810. <https://doi.org/10.1016/j.jclepro.2021.129810>
46. Friedman, J. H. (2000). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
47. Gerardi, C., Durante, M., Tufariello, M., Grieco, F., & Giovinazzo, G. (2022). Effects of time and temperature on stability of bioactive molecules, color and volatile compounds during storage of grape pomace flour. *Applied Sciences*, 12, 3956. <https://doi.org/10.3390/app12083956>
48. Gómez-Brandón, M., Martínez-Cordeiro, H., & Domínguez, J. (2021). Changes in the nutrient dynamics and microbiological properties of grape marc in a continuous-feeding vermicomposting system. *Waste Management*, 135, 1–10. <https://doi.org/10.1016/j.wasman.2021.08.004>
49. Gordillo, B., Rodríguez-Pulido, F. J., Escudero-Gilete, M. L., González-Miret, M. L., & Heredia, F. J. (2012). Comprehensive colorimetric study of anthocyanic copigmentation in model solutions: Effects of pH and molar ratio. *Journal of Agricultural and Food Chemistry*, 60(11), 2896–2905. <https://doi.org/10.1021/jf2046202>
50. Ha, V. (2023). Experimental study on remaining useful life prediction of lithium-ion batteries based on three regression models for electric vehicle applications [Preprint]. *Preprints*. <https://doi.org/10.20944/preprints202306.0999.v1>
51. Haouet, M. N., Tommasino, M., Mercuri, M. L., Benedetti, F., Bella, S. D., Framboas, M., Pelli, S., & Altissimi, M. S. (2019). Experimental accelerated shelf life determination of a ready-to-eat processed food. *Italian journal of food safety*, 7(4), 6919. <https://doi.org/10.4081/ijfs.2018.6919>
52. Hough, G., Garitta, L., & Gómez, G. (2006). Sensory shelf-life predictions by survival analysis accelerated storage models. *Food Quality and Preference*, 17(6), 468-473. <https://doi.org/10.1016/j.foodqual.2005.05.009>
53. International Organisation of Vine and Wine. (2025, April 15). *State of the global vine and wine sector* [Press conference report]. OIV. https://www.oiv.int/sites/default/files/2025-04/OIV_State_of_the_World_Vine_and_Wine_Sector_in_2024_PPT.pdf
54. Izquierdo-Bueno, I., Moraga, J., Cantoral, J. M., Carbú, M., Garrido, C., & González-Rodríguez, V. E. (2024). Smart viticulture: Applying artificial

- intelligence for improved winemaking and risk management. *Applied Sciences*, 14(22), 10277. <https://doi.org/10.3390/app142210277>
55. Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, 29(3), 31-44. <https://doi.org/10.1109/2.485891>
 56. Jalali, A., Seiedlou, S., Linke, M., & Mahajan, P. (2017). A comprehensive simulation program for modified atmosphere and humidity packaging (MAHP) of fresh fruits and vegetables. *Journal of Food Engineering*, 206, 88–97. <https://doi.org/10.1016/j.jfoodeng.2017.03.007>
 57. Jesus, M., Romani, A., Mata, F., & Domingues, L. (2022). Current Options in the Valorisation of Vine Pruning Residue for the Production of Biofuels, Biopolymers, Antioxidants, and Bio-Composites following the Concept of Biorefinery: A Review. *Polymers*, 14(9), 1640. <https://doi.org/10.3390/polym14091640>
 58. Jordão, A. M., & Cosme, F. (2021). Grape maturation. In V. K. Joshi (Ed.), *Wine production and botrytis* (Chap. 9, pp. 125–143). CRC Press. https://www.researchgate.net/publication/353637442_Grape_Maturation
 59. Jung, H. M., Lee, S., Lee, W. H., Cho, B. K., & Lee, S. H. (2018). Effect of vibration stress on quality of packaged grapes during transportation. *Engineering in Agriculture Environment & Food*, 11, 79–83. <https://doi.org/10.1016/j.eaef.2018.02.007>
 60. Kallithraka, S., Arvanitoyannis, I., El-Zajouli, A., & Kefalas, P. (2001). The application of an improved method for trans-resveratrol to determine the origin of Greek red wines. *Food Chemistry*, 75(3), 355-363. [https://doi.org/10.1016/S0308-8146\(01\)00213-8](https://doi.org/10.1016/S0308-8146(01)00213-8)
 61. Kalopesa, E., Gkrimpizis, T., Samarinas, N., Tsakiridis, N. L., & Zalidis, G. C. (2023). Rapid determination of wine grape maturity level from pH, titratable acidity, and sugar content using non-destructive in situ infrared spectroscopy and multi-head attention convolutional neural networks. *Sensors*, 23(23), 9536. <https://doi.org/10.3390/s23239536>
 62. Kalua, C. M., & Boss, P. K. (2009). Evolution of volatile compounds during the development of Cabernet Sauvignon grapes (*Vitis vinifera* L.). *Journal of Agricultural and Food Chemistry*, 57(9), 3818–3830. <https://doi.org/10.1021/jf803471n>
 63. Karastergiou, A., Gancel, A.-L., Jourdes, M., & Teissedre, P.-L. (2024). Valorization of Grape Pomace: A Review of Phenolic Composition, Bioactivity,

and Therapeutic Potential. *Antioxidants*, 13(9), 1131.
<https://doi.org/10.3390/antiox13091131>

64. Kassongo, J., Shahsavari, E., & Ball, A. S. (2020). Renewable energy from the solid-state anaerobic digestion of grape marc and cheese whey at high treatment capacity. *Biomass & Bioenergy*, 143, 105880.
<https://doi.org/10.1016/j.biombioe.2020.105880>
65. Kato-Schwartz, C. G., Corrêa, R. C. G., de Souza Lima, D., de Sá-Nakanishi, A. B., de Almeida Gonçalves, G., Seixas, F. A. V., Haminiuk, C. W. I., Barros, L., Ferreira, I. C. F. R., Bracht, A., & Peralta, R. M. (2020). Potential anti-diabetic properties of Merlot grape pomace extract: An in vitro, in silico and in vivo study of α -amylase and α -glucosidase inhibition. *Food research international (Ottawa, Ont.)*, 137, 109462. <https://doi.org/10.1016/j.foodres.2020.109462>
66. King, T., Cole, M., Farber, J. M., Eisenbrand, G., Zabaras, D., Fox, E. M., et al. (2017). Food safety for food security: Relationship between global megatrends and developments in food safety. *Trends in Food Science & Technology*, 68, 160–175.
<https://doi.org/10.1016/j.tifs.2017.08.014>
67. Koźniewski, M., Cypko, M. A., & Drużdżel, M. J. (2016). How reliable is a measure of model reliability? Bootstrap confidence intervals over validation results. In *Advances in Computer Science Research* (Vol. 13, pp. 27–41). Atlantis Press.
68. Laca, S., Gancedo, S., & Díaz, M. (2021). Assessment of the environmental impacts associated with vineyards and winemaking: A case study in mountain areas. *Environmental Science and Pollution Research*, 28, 1204–1223.
<https://doi.org/10.1007/s11356-020-10567-9>
69. Lalegani, S., Gavlighi, H. A., Azizi, M. H., & Sarteshnizi, R. A. (2018). Inhibitory activity of phenolic-rich pistachio green hull extract-enriched pasta on key type 2 diabetes relevant enzymes and glycemic index. *Food Research International*, 105, 94–101. <https://doi.org/10.1016/j.foodres.2017.11.003>
70. Li, D., Yang, Y., Chang, H., Xu, Z., & Ouyang, Q. (2025). Improved ResNet deep learning model-assisted computer vision for intelligent assessment of tencha chlorophyll content during drying. *Journal of Food Composition and Analysis*, 148(Part 4), 108545. <https://doi.org/10.1016/j.jfca.2025.108545>
71. Li, J., Tai, B. C., & Nott, D. J. (2009). Confidence interval for the bootstrap P-value and sample size calculation of the bootstrap test. *Journal of Nonparametric Statistics*, 21(5), 649-661.

72. Li Y., Chu X., Fu Z., Feng J., & Mu W. (2019). Shelf life prediction model of postharvest table grape using optimized radial basis function (RBF) neural network. *British Food Journal*, *121*(11), 2919–2936. <https://doi.org/10.1108/BFJ-03-2019-0183>
73. Lima, L., Pereira, A. I., Pintado, M., Caroch, M., & Barros, L. (2023). Multivariate Analysis in Accelerated Shelf-Life Assessment—An Overview. *Engineering Proceedings*, *56*(1), 251. <https://doi.org/10.3390/ASEC2023-15520>
74. Liu, X., Fan, P., Jiang, J., Gao, Y., Liu, C., Li, S., & Liang, Z. (2022). Evolution of volatile compounds composition during grape berry development at the germplasm level. *Scientia Horticulturae*, *293*, 110669. <https://doi.org/10.1016/j.scienta.2021.110669>
75. Lobb, K., & Chow, C. K. (2007). Fatty acid classification and nomenclature. In C. K. Chow (Ed.), *Fatty acids in foods and their health implications* (pp. 1–15). CRC Press. https://www.researchgate.net/publication/329681202_Fatty_Acid_Classification_and_Nomenclature
76. López García-Arias, Á., Okoshi, Y., Hashimoto, M., Motomura, M., & Yu, J. (2023). Recurrent residual networks contain stronger lottery tickets. *IEEE Access*, *11*, 1–1. <https://doi.org/10.1109/ACCESS.2023.3245808>
77. Lucas, M. S., Peres, J. A., & Puma, G. L. (2010). Treatment of winery wastewater by ozone-based advanced oxidation processes (O₃, O₃/UV, and O₃/UV/H₂O₂) in a pilot-scale bubble column reactor and process economics. *Separation and Purification Technology*, *72*(3), 235–241. <https://doi.org/10.1016/j.seppur.2010.01.016>
78. Machado, T.O., Portugal, I., de A C Kodel, H., Droppa-Almeida, D., dos S. Lima, M., Fathi, F., Oliveira, M.B., de Albuquerque-Júnior, R.L., Dariva, C., & Souto, E.B. (2024). Therapeutic potential of grape pomace extracts: a review of scientific evidence. *Food Bioscience*. <https://doi.org/10.1016/j.fbio.2024.104210>
79. Maesano, G., Milani, M., Nicolosi, E., D’Amico, M., & Chinnici, G. (2022). A network analysis for environmental assessment in the wine supply chain. *Agronomy*, *12*(1), 211. <https://doi.org/10.3390/agronomy12010211>
80. Maicas, S., & Mateo, J. J. (2020). Sustainability of Wine Production. *Sustainability*, *12*(2), 559. <https://doi.org/10.3390/su12020559>
81. Matrapazi, V. K., & Zabaniotou, A. (2020). Experimental and feasibility study of spent coffee grounds upscaling via pyrolysis towards proposing an eco-social

- innovation circular economy solution. *Science of the Total Environment*, 718, 137316. <https://doi.org/10.1016/j.scitotenv.2020.137316>
82. Mauro, M., Pinto, P., Settanni, L., Puccio, V., Vazzana, M., Hornsby, B. L., Fabbrizio, A., Di Stefano, V., Barone, G., & Arizza, V. (2022). Chitosan Film Functionalized with Grape Seed Oil—Preliminary Evaluation of Antimicrobial Activity. *Sustainability*, 14(9), 5410. <https://doi.org/10.3390/su14095410>
83. Mistry, I., Tanwar, S., Tyagi, S., & Kumar, N. (2020). Blockchain for 5G-enabled IoT for industrial automation: A systematic review, solutions, and challenges. *Mechanical Systems and Signal Processing*, 135, 106382. <https://doi.org/10.1016/j.ymsp.2019.106382>
84. Moggia, C., Graell, J., Lara, I., Gonzalez, G., & Lobos, G. A. (2017). Firmness at harvest impacts postharvest fruit softening and internal browning development in mechanically damaged and non-damaged highbush blueberries (*Vaccinium corymbosum* L.). *Frontiers in Plant Science*, 8, 535. <https://doi.org/10.3389/fpls.2017.00535>
85. Montealegre, R. R., Peces, R. R., Vozmediano, J. C., Gascueña, J. M., & Romero, E. G. (2006). Phenolic compounds in skins and seeds of ten grape *Vitis vinifera* varieties grown in a warm climate. *Journal of Food Composition and Analysis*, 19(6-7), 687-693. <https://doi.org/10.1016/j.jfca.2005.05.003>
86. Muhlack, R. A., Potumarthi, R., & Jeffery, D. W. (2018). Sustainable wineries through waste valorisation: A review of grape marc utilisation for value-added products. *Waste Management*, 72, 99–118. <https://doi.org/10.1016/j.wasman.2017.11.011>
87. Muñoz-Bernal, Ó. A., Coria-Oliveros, A. J., de la Rosa, L. A., Rodrigo-García, J., del Rocío Martínez-Ruiz, N., Sayago-Ayerdi, S. G., & Alvarez-Parrilla, E. (2021). Cardioprotective effect of red wine and grape pomace. *Food Research International*, 140, 110069. <https://doi.org/10.1016/j.foodres.2020.110069>
88. Nakatsu, R. (2023). Validation of machine learning ridge regression models using Monte Carlo, bootstrap, and variations in cross-validation. *Journal of Intelligent Systems*, 32(1), 20220224 <https://doi.org/10.1515/jisys-2022-0224>
89. Namazi, P., Alizadeh Behbahani, B., Noshad, M., Vasiee, A., Taki, M., & Jooyandeh, H. (2025). Active coating of *Lepidium perfoliatum* mucilage and *Lactiplantibacillus plantarum* OC-D: Shelf-life extension and support vector regression prediction. *Applied Food Research*, 5(2), 101405. <https://doi.org/10.1016/j.afres.2025.101405>

90. Navarro, P. J., Miller, L., Díaz-Galián, M. V., Gila-Navarro, A., Aguila, D. J., & Egea-Cortines, M. (2022). A novel ground truth multispectral image dataset with weight, anthocyanins, and Brix index measures of grape berries tested for its utility in machine learning pipelines. *GigaScience*, *11*, giac052. <https://doi.org/10.1093/gigascience/giac052>
91. Nirmal, N., Mahale, K. R., Rathod, N. B., Siddiqui, S. A., & Dhar, B. K. (2025). Winery waste: A sustainable approach for bioactive compound extraction and various industrial applications. *Process Safety and Environmental Protection*, *193*, 760–771. <https://doi.org/10.1016/j.psep.2024.11.106>
92. Oancea, S. (2021). A review of the current knowledge of thermal stability of anthocyanins and approaches to their stabilization to heat. *Antioxidants*, *10*(9), 1337. <https://doi.org/10.3390/antiox10091337>
93. Oliveira, H. M., Tugnolo, A., Fontes, N., Marques, C., Geraldés, Á., Jenne, S., Zappe, H., Graça, A., Giovenzana, V., Beghi, R., Guidetti, R., Piteira, J., & Freitas, P. (2024). An autonomous Internet of Things spectral sensing system for in-situ optical monitoring of grape ripening: Design, characterization, and operation. *Computers and Electronics in Agriculture*, *217*, 108599. <https://doi.org/10.1016/j.compag.2023.108599>
94. Oren-Shamir, M. (2009). Does anthocyanin degradation play a significant role in determining pigment concentration in plants?. *Plant Science*, *177*(4), 310-316. <https://doi.org/10.1016/j.plantsci.2009.06.015>
95. Pang, B., Nijkamp, E., & Wu, Y. N. (2019). Deep Learning With TensorFlow: A Review. *Journal of Educational and Behavioral Statistics*, *45*(2), 227-248. <https://doi.org/10.3102/1076998619872761>
96. Panza, O., Alfieri, M. L., Del Nobile, M. A., Mugnaini, G., Conte, A., Panzella, L., & Bonini, M. (2025). Effects of dehydration temperature on physico-chemical, antioxidant and antimicrobial properties of grape pomace powder. *LWT*, *227*, 118029. <https://doi.org/10.1016/j.lwt.2025.118029>
97. Patras, A., Brunton, N. P., O'Donnell, C., & Tiwari, B. K. (2010). Effect of thermal processing on anthocyanin stability in foods: Mechanisms and kinetics of degradation. *Trends in Food Science & Technology*, *21*(1), 3–11. <https://doi.org/10.1016/j.tifs.2009.07.004>
98. Peixoto, C. M., Dias, M. I., Alves, M. J., Calhelha, R. C., Barros, L., Pinho, S. P., & Ferreira, I. C. F. R. (2018). Grape pomace as a source of phenolic compounds

- and diverse bioactive properties. *Food Chemistry*, 253, 132–138. <https://doi.org/10.1016/j.foodchem.2018.01.163>
99. Perra, M., Bacchetta, G., Muntoni, A., De Gioannis, G., Castangia, I., Rajha, H. N., Manca, M. L., & Manconi, M. (2022). An outlook on modern and sustainable approaches to the management of grape pomace by integrating green processes, biotechnologies and advanced biomedical approaches. *Journal of Functional Foods*, 98, 105276. <https://doi.org/10.1016/j.jff.2022.10527>
100. Piombino, P., Pittari, E., Genovese, A., Bellincontro, A., & Moio, L. (2025). Postharvest dehydration of red grapes: impact of temperature and water-loss conditions on free and glycosylated volatile metabolites of exocarp and epicarp of Nebbiolo and Aleatico varieties. *Journal of the science of food and agriculture*, 105(1), 123–131. <https://doi.org/10.1002/jsfa.13810>
101. Rahim, M. A., Rahman, M., Islam, M. S., Joseph, C. I., Rahman, M. A., Ramasamy, D., Rahman, M., & Islam, M. Z. (2025). Attentive neural networks and meta-learning integration for revolutionary vehicular engine health monitoring. *Mechanical Systems and Signal Processing*, 238, 113141. <https://doi.org/10.1016/j.ymsp.2025.113141>
102. Ramos, M. C., Ibáñez Jara, M. Á., Rosillo, L., & Salinas, M. R. (2024). Effect of temperature and water availability on grape phenolic compounds and their extractability in Merlot grown in a warm area. *Scientia Horticulturae*, 337, 113475. <https://doi.org/10.1016/j.scienta.2024.113475>
103. Rashvand, M., Ren, Y., Sun, D.-W., Senge, J., Krupitzer, C., Fadiji, T., Sanzo Miró, M., Shenfield, A., Watson, N. J., & Zhang, H. (2025). Artificial intelligence for prediction of shelf-life of various food products: Recent advances and ongoing challenges. *Trends in Food Science & Technology*, 159, 104989. <https://doi.org/10.1016/j.tifs.2025.104989>
104. Ribeiro, L.F., Ribani, R.H., Francisco, T.M., Soares, A.A., Pontarolo, R., & Haminiuk, C.W. (2015). Profile of bioactive compounds from grape pomace (*Vitis vinifera* and *Vitis labrusca*) by spectrophotometric, chromatographic and spectral analyses. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences*, 1007, 72-80. <https://doi.org/10.1016/j.jchromb.2015.11.005>
105. Rodríguez-Pulido, F. J., Ferrer-Gallego, R., González-Miret, M. L., Rivas-Gonzalo, J. C., Escribano-Bailón, M. T., & Heredia, F. J. (2012a). Preliminary study to determine the phenolic maturity stage of grape seeds by computer vision. *Analytica Chimica Acta*, 732, 78–82. <https://doi.org/10.1016/j.aca.2012.01.005>

106. Rodríguez-Pulido, F. J., Gómez-Robledo, L., Melgosa, M., Gordillo, B., González-Miret, M. L., & Heredia, F. J. (2012b). Ripeness estimation of grape berries and seeds by image analysis. *Computers and Electronics in Agriculture*, *82*, 128–133. <https://doi.org/10.1016/j.compag.2012.01.004>
107. Rodríguez-Pulido, F. J., Mora-Garrido, A. B., González-Miret, M. L., & Heredia, F. J. (2022). Research Progress in Imaging Technology for Assessing Quality in Wine Grapes and Seeds. *Foods*, *11*(3), 254. <https://doi.org/10.3390/foods11030254>
108. Rohani, A., Taki, M., & Abdollahpour, M. (2018). A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part I). *Renewable Energy*, *115*, 411–422. <https://doi.org/10.1016/j.renene.2017.08.061>
109. Rousseau, F., Drumetz, L., & Fablet, R. (2020). Residual networks as flows of diffeomorphisms. *Journal of Mathematical Imaging and Vision*, *62*(3), 365–375. <https://doi.org/10.1007/s10851-019-00890-3>
110. Salman, H. A., Kalakech, A., & Steiti, A. (2024). Random forest algorithm overview. *Babylonian Journal of Machine Learning*, *2024*, 69-79. <https://doi.org/10.58496/BJML/2024/007>
111. Samarinas, N., Tsakiridis, N., Tziolas, N., Kalopesa, E., Karyotis, K., Bochtis, D., Zalidis, G., & Doulgeri, Z. (2020). Non-destructive estimation of grape ripeness in Syrah variety via VNIR–SWIR spectroscopy. *Proceedings*, *4*(1), 1–5.
112. Sanmartin, C., Modesti, M., Venturi, F., Brizzolara, S., Mencarelli, F., & Bellincontro, A. (2021). Postharvest water loss of wine grape: When, what and why. *Metabolites*, *11*, 318. <https://doi.org/10.3390/metabo11050318>
113. Santillán, D., Iglesias, A., La Jeunesse, I., Garrote, L., & Sotes, V. (2019). Vineyards in transition: A global assessment of the adaptation needs of grape producing regions under climate change. *Science of the Total Environment*, *657*, 839–852. <https://doi.org/10.1016/j.scitotenv.2018.12.079>
114. Savio, M., Ferraro, D., Maccario, C., Vaccarone, R., Jensen, L. D., Corana, F., Mannucci, B., Bianchi, L., Cao, Y., & Stivala, L. A. (2016). Resveratrol analogue 4,4'-dihydroxy-trans-stilbene potently inhibits cancer invasion and metastasis. *Scientific Reports*, *6*, 19973. <https://doi.org/10.1038/srep19973>
115. Savio, M., Ferraresi, A., Corpina, C., Vandenberghe, S., Scarlata, C., Sottile, V., Morini, L., Garavaglia, B., Isidoro, C., & Stivala, L. A. (2022). Resveratrol and its analogue 4,4'-dihydroxy-trans-stilbene inhibit Lewis Lung Carcinoma growth in

- vivo through apoptosis, autophagy and modulation of the tumour microenvironment in a murine model. *Biomedicines*, *10*(8), 1784. <https://doi.org/10.3390/biomedicines10081784>
116. Sha, X., Liu, Y., Chen, H., Zhang, Z., & Ge, M. (2021). Anthocyanin degradation and the underlying molecular mechanism in a red-fleshed grape variety. *LWT*, *151*, 112198. <https://doi.org/10.1016/j.lwt.2021.112198>.
 117. Shehzad, K. (2025). Predictive AI models for food spoilage and shelf-life estimation. *Global Trends in Science and Technology*, *1*(1), 75–94. <https://doi.org/10.70445/gtst.1.1.2025.75-9475>
 118. Singh, U., Rizwan, M., Alaraj, M., & Alsaidan, I. (2021). A machine learning-based Gradient Boosting Regression approach for wind power production forecasting: a step towards smart grid environments. *Energies*, *14*(16), 5196. <https://doi.org/10.3390/en14165196>
 119. Sirohi, R., Tarafdar, A., Singh, S., Negi, T., Gaur, V. K., Gnansounou, E., & Bharathiraja, B. (2020). Green processing and biotechnological potential of grape pomace: Current trends and opportunities for sustainable biorefinery. *Bioresource Technology*, *314*, 123771. <https://doi.org/10.1016/j.biortech.2020.123771>
 120. Smalheiser, N. R. (2017). Correlation and other concepts you should know. In N. R. Smalheiser (Ed.), *Data Literacy* (pp. 169–185). Academic Press. <https://doi.org/10.1016/B978-0-12-811306-6.00013-0>
 121. Solfrizzo, M., Panzarini, G., & Visconti, A. (2008). Determination of ochratoxin A in grapes, dried vine fruits, and winery byproducts by high-performance liquid chromatography with fluorometric detection (HPLC–FLD) and immunoaffinity cleanup. *Journal of Agricultural and Food Chemistry*, *56*, 11081–11086. <https://doi.org/10.1021/jf802380d>
 122. Sousa Gallagher, M. J., & Mahajan, P. V. (2011). The stability and shelf life of fruit and vegetables. In D. Kilcast & P. Subramaniam (Eds.), *Food and beverage stability and shelf life* (pp. 641–656). Woodhead Publishing. <https://doi.org/10.1533/9780857092540.3.641>
 123. Sumayli, A. (2023). Development of advanced machine learning models for optimization of methyl ester biofuel production from papaya oil: Gaussian process regression (GPR), multilayer perceptron (MLP), and K-nearest neighbor (KNN) regression models. *Arabian Journal of Chemistry*, *16*(7), 104833. <https://doi.org/10.1016/j.arabjc.2023.104833>

124. Tchatchoua, P., Graton, G., Ouladsine, M., & Christaud, J.-F. (2023). Application of 1D ResNet for multivariate fault detection on semiconductor manufacturing equipment. *Sensors*, 23(22), 9099. <https://doi.org/10.3390/s23229099>
125. Torres-Sánchez, R., Martínez-Zafra, M. T., Castillejo, N., Guillamón-Frutos, A., & Artés-Hernández, F. (2020). Real-Time Monitoring System for Shelf Life Estimation of Fruit and Vegetables. *Sensors*, 20(7), 1860. <https://doi.org/10.3390/s20071860>
126. Tsakiridis, N. L., Samarinas, N., Kokkas, S., Kalopesa, E., Tziolas, N. V., & Zalidis, G. C. (2023). In situ grape ripeness estimation via hyperspectral imaging and deep autoencoders. *Computers and Electronics in Agriculture*, 212, 108098. <https://doi.org/10.1016/j.compag.2023.108098>
127. Ventura, V. (2010). Bootstrap tests of hypotheses. In S. Grün & S. Rotter (Eds.), *Analysis of parallel spike trains* (Springer Series in Computational Neuroscience, Vol. 7). Springer. https://doi.org/10.1007/978-1-4419-5675-0_18
128. Verzeletti, T.N., Priamo, W.L. & Flores, S.S. Use of information technology for traceability and quality control: case study in a brazilian winery. *BIO Web Conf.* 2023, 68, 04018. <https://doi.org/10.1051/bioconf/20236804018>
129. Viveros Santos, I., Renaud-Gentié, C., Roux, P., Levasseur, A., Bulle, C., Deschênes, L., & Boulay, A.-M. (2023). Prospective life cycle assessment of viticulture under climate change scenarios: Application on two case studies in France. *Science of the Total Environment*, 880, 163288. <https://doi.org/10.1016/j.scitotenv.2023.163288>
130. VEGEA S.r.l. (n.d.). <https://www.vegeacompany.com/>
131. Wang, C., You, Y., Huang, W., & Zhan, J. (2024). The high-value and sustainable utilization of grape pomace: A review. *Food Chemistry: X*, 24, 101845. <https://doi.org/10.1016/j.fochx.2024.101845>
132. Welke, J. E. (2019). Fungal and mycotoxin problems in grape juice and wine industries. *Current Opinion in Food Science*, 29, 7–13. <https://doi.org/10.1016/j.cofs.2019.06.009>
133. Xu, M., Sun, J., Cheng, J., Yao, K., Shi, L., & Zhou, X. (2024). Non-destructive estimation for Kyoho grape shelf-life using Vis/NIR hyperspectral imaging and deep learning algorithm. *Infrared Physics & Technology*, 142, 105532. <https://doi.org/10.1016/j.infrared.2024.105532>

134. Yu, J., & Pedroso, I. R. (2023). Mycotoxins in cereal-based products and their impacts on the health of humans, livestock animals and pets. *Toxins*, 15(8), 480. <https://doi.org/10.3390/toxins15080480>
135. Yu, J., Smith, I., Karlton-Senaye, B., & Mikiashvili, N. (2018). Impacts of different drying methods on mold viability and ochratoxin A content of grape pomace. *International Journal of Applied Agricultural Sciences*, 4(2), 2. <https://doi.org/10.11648/j.ijaas.20180402.12>
136. Yu, J., Smith, I. N., & Mikiashvili, N. (2020). Reducing ochratoxin A content in grape pomace by different methods. *Toxins*, 12(7), 424. <https://doi.org/10.3390/toxins12070424>
137. Yu, J., Tseng, Y., Pham, K., Liu, M., & Beckles, D. M. (2022). *Starch and sugars as determinants of postharvest shelf life and quality: Some new and surprising roles*. *Current Opinion in Biotechnology*, 78, 102844. <https://doi.org/10.1016/j.copbio.2022.102844>
138. Yu, J., Wang, M., Li, Z., Tchuenbou-Magaia, F., Wani, A. A., Zhu, P., Fadiji, T., & Liu, Y. (2024). Preserving freshness: Innovations for fresh-eating fruit distribution and damage prevention – A review. *Food Packaging and Shelf Life*, 44, 101323. <https://doi.org/10.1016/j.fpsl.2024.101323>
139. Zuñiga, A., Mora, M., Oyarce, M., & Fredes, C. (2014). Grape maturity estimation based on seed images and neural networks. *Engineering Applications of Artificial Intelligence*, 35, 95–104. <https://doi.org/10.1016/j.engappai.2014.06.007>

RINGRAZIAMENTI

Desidero innanzitutto esprimere sincera gratitudine alla mia Relatrice, la Professoressa Savio, per il supporto e la disponibilità dimostrati durante tutto il percorso di tesi. Sono sempre stata fiera e pienamente soddisfatta della mia scelta.

Un ringraziamento speciale va alla mia famiglia e, in particolare, ai miei genitori, che mi sono stati accanto con costanza anche nei momenti in cui ho intrapreso scelte non sempre semplici, ma di cui, in fondo, non mi pento.

Ringrazio tutti coloro che mi hanno accompagnata lungo questo percorso, anche nelle sue deviazioni, e in particolare chi è sempre stato presente.

Un pensiero sincero va alle preziose amicizie conosciute a Pavia: senza saperlo, mi hanno regalato la possibilità di affrontare la vita con leggerezza, di esprimermi in modo autentico e di crescere. Sono diventata la persona che sono oggi anche grazie a loro; sono state il mio porto sicuro lontano da casa, anche quando le nostre strade hanno iniziato a dividersi. Separarmi da loro sarà difficile, ma sono consapevole che è una tappa naturale da percorrere. Porterò con me ciò che mi hanno dato, continuando a cercare legami che mi facciano sentire altrettanto fortunata.

Un pensiero grato va a chi si è allontanato lungo il cammino: anche da queste esperienze ho imparato che non tutto il male viene per nuocere.

Ringrazio, infine, chi ha messo nuvole lungo il mio percorso e chi mi ha travolta in cicloni imprevedibili: mi ha insegnato a trovare la mia direzione anche senza sole e a superare i miei limiti. Questo lavoro, insieme a tutto ciò che lo ha accompagnato, ne è testimonianza.

Questo capitolo è stato ricco di svolte inaspettate, ma ogni esperienza ha un suo significato e avere fiducia nel processo è ciò che permette di avanzare.

È con questa consapevolezza che guardo al mio futuro.

Annalisa Simone