



UNIVERSITÀ
DI PAVIA

Dipartimento di Scienze Economiche e Aziendali Corso di

Laurea Magistrale in International Business and

Entrepreneurship

Enhancing AI Data Comprehension through Data Visualization

Relatore:

Prof. Bartosiak Marcin

Tesi di Laurea
di Ricci Renato

Matr. n. 501963

Anno Accademico 2023-2024

Acknowledgement

As a support for content creation and clarity expression, generative AI tools have been used, such as ChatGPT.

Contents

| | |
|--|----|
| Acknowledgement | 2 |
| Contents | 3 |
| Introduction | 4 |
| What is data visualization | 6 |
| Definition and its significance in data analysis and communication | 6 |
| Main tools of Data Viz..... | 7 |
| What is Artificial Intelligence..... | 11 |
| History of AI | 11 |
| Type of data | 12 |
| AI technologies: Machine learning, deep learning and natural language processing | 14 |
| AI applications..... | 17 |
| Ethical consideration on AI..... | 22 |
| Challenges of interpreting and understanding this data (data transparency) | 25 |
| Advantages of using data visualization to communicate AI-generated insights..... | 27 |
| Better understanding of data through visualization tools | 28 |
| Facilitate decision making..... | 31 |
| Case Studies..... | 33 |
| Challenging and limitations of data visualization in the context of AI-generated data | 36 |
| Handling High-Dimensional Data..... | 37 |
| Scalability Issues: challenges in scaling visualizations as data volume and real-time data streams increase..... | 39 |
| Future research | 41 |
| Bibliography | 42 |

Introduction

In today's data-driven world, the ability to analyze and interpret vast amounts of data has become a critical asset for organizations and individuals alike. As artificial intelligence (AI) continues to evolve and integrate into various sectors, the volume and complexity of data generated by these technologies have exponentially increased. However, the size and complexity of AI-generated data create significant challenges for end-users in understanding and making decisions.

The focus of this topic lies in the need to find a solution to understand data generated by Artificial Intelligence. Typically, this kind of data is hard to comprehend and especially difficult to interpret. Data visualization appears to be a promising solution to achieve this goal. In this thesis, I will focus on the types of data that AI can produce, the industries in which AI is most beneficial, and how data visualization can improve the interpretability of this data. The purpose of this thesis is not to delve deeply into the technical details of AI and data generation. Instead, it will provide a general overview of the main AI topics. The literature on this subject is still incomplete, and much research is needed to fully explain and prove the benefits that data visualization can provide for AI-generated data. A deep dive into this topic is essential to offer companies new ways to handle and fully exploit the vast amounts of data generated by their activities. Finally, the goal of this thesis is to synthesize the benefits of data visualization for interpreting AI-generated data based on current literature. The first section of this thesis provides a comprehensive overview of data visualization, detailing various techniques and tools, with a particular focus on Tableau, one of the leading data visualization platforms. It also delves into the fundamentals of artificial intelligence, outlining the types of data generated and the inherent challenges in interpreting this data.

The second section highlights the advantages of using data visualization to communicate AI-generated insights, emphasizing how visualization tools can improve data comprehension, support decision-making, and enhance user engagement. This section also includes case studies to illustrate practical applications and benefits.

The final section addresses the challenges and limitations of data visualization in the context of AI-generated data, exploring areas for future research and potential advancements.

What is data visualization

Definition and its significance in data analysis and communication

Very explanatory is this definition by R. S. Raghav saying that “The data visualization provides the user to acquire more knowledge about the raw data which is collected from the variety of sources. The visualization can be done by using the dashboards, where the undetected text, patterns and correlations can be easily visualized by using the visualization software.”¹

“Data visualization is the representation of information and data using charts, graphs, maps, and other visual tools. These visualizations allow us to easily understand any patterns, trends, or outliers in a data set.”²

These tools also help people with lack of knowledge on the topic to understand data. In fact, we will be analyzing how data visualization helps users to understand data coming from AI.

The history of data visualization goes along with human civilization, but the first representation of statistical data comes from 1644 when Michael Florent van Langren, a Dutch astronomer who published maps illustrating his calculations regarding eclipses.

With the introduction of computers, data visualization had a huge boost. Microsoft Excel set the path of data visualization that we know nowadays.

Data visualization was and still is used to understand large amount of data. Comprehending the proportion of a total is much easier using a pie chart for example than writing down numbers in a table.

¹ R. S. Raghav, S. Pothula, T. Vengattaraman and D. Ponnurangam, "A survey of data visualization tools for analyzing large volume of data in big data platform," 2016 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2016, pp. 1-6, doi: 10.1109/CESYS.2016.7889976. keywords: {Data visualization;Companies;Big Data;Social network services;Data analysis;Data Visualization;Big Data;SNA;BDA;BIA},

² <https://www.coursera.org/articles/data-visualization>

Data visualization is widely used to grasp insights from data in relatively small amounts of time. Consequently, the decision making process is reduced in time. This is crucial for many industries, for companies that have to steer in some directions or resolve inefficiencies in very limited time. In terms of communication, data visualization plays a crucial role in bridging the gap between technical experts and non-technical audiences. It transforms complex analytical outputs into visually accessible narratives, fostering better communication and collaboration among diverse stakeholders. By making data more comprehensible and engaging, data visualization enhances the impact of insights and supports data-driven decision-making at all levels of an organization.

Main tools of Data Viz

There are many tools to create dashboards to better analyze data. They differ in terms of functionalities, complexity and user-friendliness. The main software producers (and the ones we will focus on are Salesforce and Microsoft). Salesforce owns Tableau, a leading data visualization and business intelligence platform known for its ease of use, flexibility, and robust capabilities, Tableau has become a popular choice among businesses, analysts, and data scientists for creating interactive dashboards, reports, and visualizations. One of Tableau's defining characteristics is its intuitive drag-and-drop interface. This design allows users to create complex visualizations without extensive coding or technical expertise. Users can drag fields from their data sources into Tableau's workspace to create various visualizations, such as charts, graphs, maps, and dashboards. This simplicity makes Tableau accessible to a wide range of users, from business analysts to data scientists.

Moreover, Salesforce's product has the ability to create interactive dashboards. Users can combine multiple visualizations into a single dashboard and enable interactive features like filters, drill-downs, and actions. This interactivity allows end-users to explore data dynamically, gaining deeper insights and answering specific questions on the fly.

Regarding how the software connects to data, Tableau supports a wide range of data sources, allowing users to connect to databases, cloud services, spreadsheets, and more. It integrates with popular platforms like Excel, SQL databases, Google Analytics, Salesforce, and AWS. This broad connectivity enables users to work with data from various sources, facilitating comprehensive analysis and visualization.

Tableau Server provides a centralized platform where users can publish, share, and interact with Tableau visualizations and dashboards. This centralization ensures that all users have access to the latest reports and insights, fostering collaboration and a unified approach to data analysis.

It offers robust security features to ensure that only authorized users can access sensitive data. Administrators can define user roles and permissions, controlling who can view, edit, or publish content. Role-based security and integration with authentication systems (like Active Directory or OAuth) provide a secure environment for data sharing.

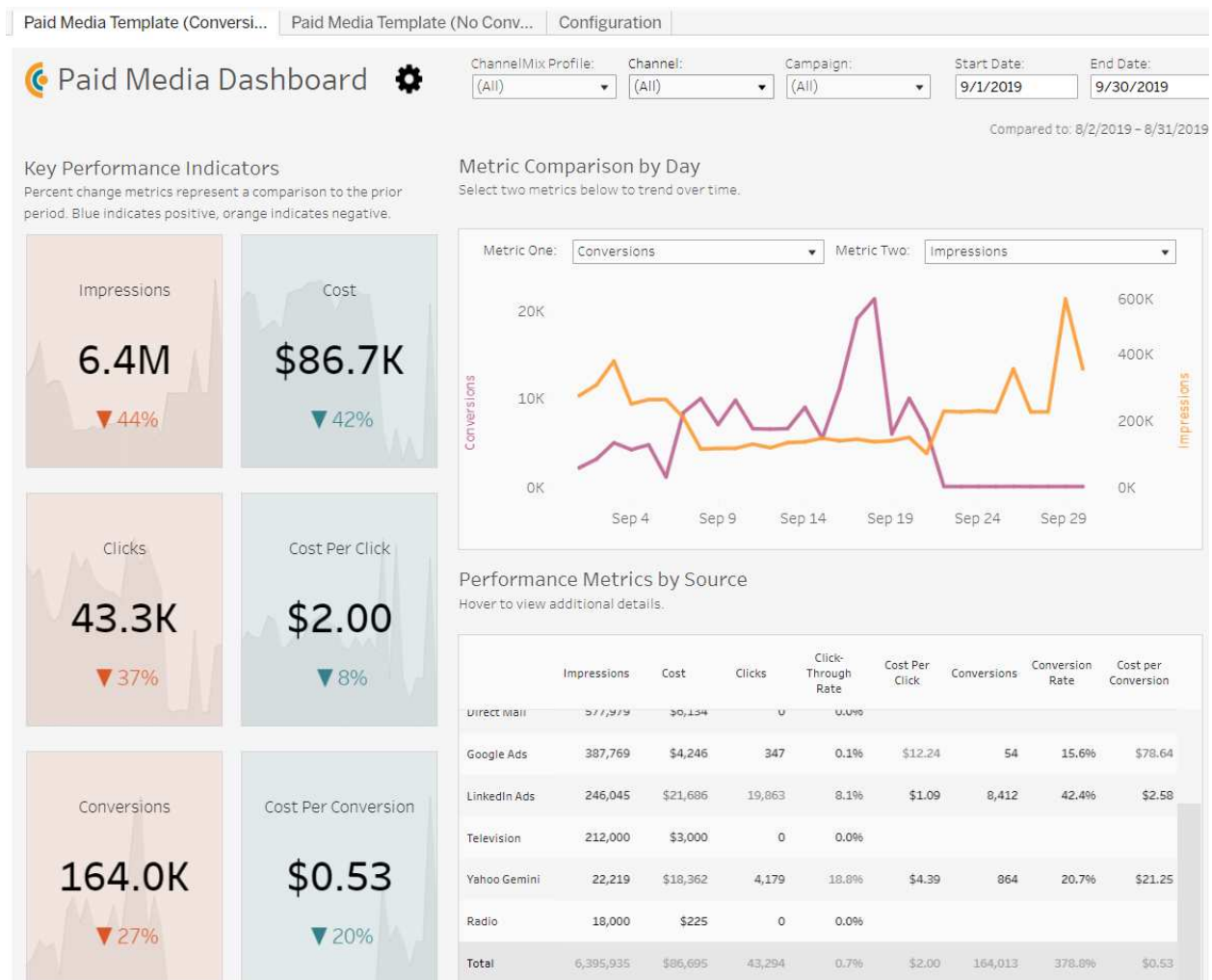
Workbooks' data sources (where multiple dashboards containing multiple graphs) can be updated on schedule. This automation ensures that dashboards and visualizations are always up-to-date, providing users with the latest information. Scheduled tasks can run at specific times or frequencies, allowing organizations to maintain a consistent reporting cadence.

Power BI

Power BI is a collection of software services, apps, and connectors that can create insights from unrelated sources of data.

The biggest advantage of this tool is the easy connection to the Microsoft tools.

Power BI has a Windows desktop application called Power BI Desktop; an online software as a service (SaaS) service called the Power BI service; Power BI Mobile apps for Windows, iOS, and Android devices. The connection among these 3 services make the user able to create, explore and share business insights.



This is an example of a Tableau dashboard showing insights about paid media. We can see how easy to understand and how well it communicates the numbers regarding total impressions, cost, clicks etc. The use of color is essential to communicate even more efficiently to the audience, in fact red is used for a decrease compared to the previous period, and blue to highlight a growth. To the right side we go through a dual-axis chart, in which we see two parameters in the two y-axis. The parameters can be changed by the user and see how they performed over time. Finally, in the bottom right is shown a table. On top of the dashboard we can see the filters that the user can change to display the desired data.

This is just an example of how easy and efficient data can be communicated through a data visualization tool.

What is Artificial Intelligence

History of AI

AI can be defined as “[..]the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings”³, as stated by Copeland.

The concept of AI is based on the idea of building machines capable of performing tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and language translation.

Muthukrishnan in his research “Brief history of artificial intelligence” (2020) describes “ the journey of AI began in the mid-20th century, rooted in classical philosophers' attempts to describe human thinking as a symbolic system.” The formal founding of AI as an academic discipline is often attributed to a workshop at Dartmouth College in 1956, where the term "Artificial Intelligence" was first coined by John McCarthy. Here, the vision was set to explore the possibilities of machines performing functions associated with human intelligence.

Throughout the decades, AI has seen cycles of significant advances, followed by setbacks and disappointments – often referred to as the "AI winter" – due to the overestimation of AI capabilities and underestimation of the complexity of human cognition. However, the resurgence of interest in AI in the 21st century has been fueled by the availability of large amounts of data (big data), advances in machine learning algorithms, and improvements in computing power.

³ Copeland, B.J.. "artificial intelligence". *Encyclopedia Britannica*, 18 Jun. 2024, <https://www.britannica.com/technology/artificial-intelligence>. Accessed 20 June 2024.

Type of data

Summarizing what Muthukrishnan declared in his research, AI systems process both structured data, such as tables and databases, and unstructured data, including text, images, and videos. Structured data is straightforward to analyze and is typically used in clearly defined problem spaces, whereas unstructured data requires more complex approaches like deep learning to parse and understand.

AI systems are prolific data generators, creating diverse datasets that are instrumental in enhancing machine learning models, providing insights, or crafting new content. These types of data vary in form and function, addressing needs across various industries and applications.

One of the primary outputs of AI is predictive data. This involves AI systems making forecasts or predictions based on historical data. For example, in the business sector, AI might predict future sales trends, while in finance, it could assess risks or predict stock performances. Healthcare applications see AI predicting patient disease risks or potential treatment outcomes, showcasing the model's ability to extrapolate from existing datasets to future probabilities.

AI is also adept at generating synthetic data, which serves as an alternative to real data and is particularly valuable where real data is sensitive or scarce. For instance, Generative Adversarial Networks (GANs) can produce images for training computer vision systems, ensuring privacy and scalability. Similarly, AI-generated text helps in training natural language processing models without the need for extensive real-world text corpuses.

Anomaly data is another type generated by AI, where the system identifies outliers or unusual patterns critical for applications like fraud detection in financial transactions or network security monitoring. Such capabilities enable proactive measures against potential threats and irregularities.

In fields like robotics and autonomous vehicles, simulation data created by AI plays a critical role. These data help machines to understand and navigate complex real-world environments through virtual simulations, enhancing learning and operational efficiency without the risks associated with real-world testing.

AI systems enhance user experience by generating user interaction data. This data helps in personalizing services, as seen in recommendation systems that suggest products or content based on user behavior analysis. Adaptive learning systems in education utilize this data type to customize educational content, catering to individual learning paces and preferences.

Moreover, AI can produce metadata, which provides insights about other data, facilitating easier management and analysis. This includes automatically generating tags or captions for images and categorizing content like news articles or social media posts.

Lastly, AI generates feedback data through interactive applications such as chatbots and virtual assistants. These systems evolve with each user interaction, continuously improving to offer more accurate and satisfactory responses. Similarly, user engagement metrics generated by AI can help optimize the design of apps and websites by identifying features or content that engage users effectively.

AI technologies: Machine learning, deep learning and natural language processing

Machine Learning

Machine Learning is a “[...] branch of computational algorithms that are designed to emulate human intelligence by learning from the surrounding environment.”⁴ More specifically, it is a subset of AI focused on developing algorithms that allow computers to learn patterns and make decisions from data without explicit programming. It involves training models on labeled datasets to recognize patterns and make predictions or classifications. Common techniques in machine learning include supervised learning, unsupervised learning, and reinforcement learning. Applications range from predictive analytics and recommendation systems to image and speech recognition.

Deep Learning

“Deep learning is a subset of machine learning that uses multilayered neural networks, called deep neural networks, to simulate the complex decision-making power of the human brain.”⁵ Deep learning algorithms, inspired by the structure and function of the human brain, excel at tasks such as image recognition, natural language processing, and speech recognition. Notable architectures include convolutional neural networks (CNNs) for image analysis and recurrent neural networks (RNNs) for sequential data processing.

Generative AI

⁴ El Naqa, I., Murphy, M.J. (2015). What Is Machine Learning?. In: El Naqa, I., Li, R., Murphy, M. (eds) Machine Learning in Radiation Oncology. Springer, Cham. https://doi.org/10.1007/978-3-319-18305-3_1

⁵ <https://www.ibm.com/topics/deep-learning>

Generative AI involves algorithms that can generate new, synthetic data resembling existing datasets. Generative models, such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), learn to generate realistic data samples by capturing underlying patterns and distributions in the training data. In detail, “a generative AI model refers to generative modeling that is instantiated with a machine learning architecture (e.g., a deep neural network) and, therefore, can create new data samples based on learned patterns.”⁶

Generative AI finds applications in image generation, text synthesis, and creative tasks such as artwork and music composition.

Speech Recognition

Speech recognition, also known as automatic speech recognition (ASR), is a subfield of AI focused on converting spoken language into text. ASR systems use acoustic and language models to analyze audio input and transcribe it into text. Recent advancements in deep learning, particularly with recurrent neural networks and attention mechanisms, have significantly improved the accuracy and performance of speech recognition systems. Applications include virtual assistants, voice-controlled devices, and transcription services.

Computer Vision

In recent years, computer vision has emerged as a pivotal component of artificial intelligence, enabling machines to interpret and understand visual information much like humans do. This interdisciplinary field focuses on developing algorithms and systems that can analyze images and videos, extract meaningful insights, and make decisions based on visual input. Through techniques

⁶ Feuerriegel, S., Hartmann, J., Janiesch, C. et al. Generative AI. *Bus Inf Syst Eng* 66, 111–126 (2024).

such as image processing, feature extraction, object detection, and semantic segmentation, computer vision algorithms can identify objects, recognize patterns, and comprehend complex scenes. Applications of computer vision span various domains, including autonomous vehicles, surveillance systems, medical imaging, augmented reality, and retail. As advancements in deep learning continue to drive progress in computer vision, the technology's potential for transforming industries and enhancing human-computer interaction remains profound.

AI applications

To list AI applications to different industries, I followed what Malik stated in his paper.⁷

AI has permeated various industries, revolutionizing traditional practices and enabling new capabilities that were previously unimaginable. These are the main industries where AI is an disruptive factor:

1. Healthcare

AI is transforming healthcare by improving diagnostics, personalizing treatment, and managing patient data. AI algorithms can analyze complex medical data, such as imaging scans, to detect anomalies like tumors or fractures more accurately and faster than human practitioners. AI-powered systems like IBM Watson are being used for oncology diagnostics by Pfizer, offering treatment options by analyzing the medical literature and patient records. Additionally, AI is used in predictive analytics to forecast patient outcomes, optimize hospital operations, and manage chronic diseases by monitoring patient data in real time.

Finally, AI has the potential to shorten clinical trial cycle durations while also enhancing productivity and clinical development outcomes.

2. Finance:

The application of AI in finance encompasses a variety of innovative solutions that enhance operational efficiency, customer experience, and strategic decision-making. In particular:

⁷ Shiza Malik, Khalid Muhammad, Yasir Waheed, Artificial intelligence and industrial applications-A revolution in modern industries, Ain Shams Engineering Journal,2024

Algorithmic Trading

AI significantly transforms the landscape of trading and investment management by automating complex, quantitative trading strategies. Using machine learning algorithms, AI systems can analyze large datasets quickly, identify patterns, and execute trades at optimal times that humans cannot match for speed or efficiency. These algorithms consider variables such as stock price movements, economic indicators, and consumer behavior, enabling them to adapt to new information as it becomes available. This agility helps in maximizing returns and reducing risks, particularly in high-frequency trading.

Fraud Detection and Risk Management

AI enhances the ability of financial institutions to detect and prevent fraudulent activities by analyzing behavioral patterns and spotting anomalies that might indicate fraudulent behavior. Machine learning models are trained on historical transaction data, enabling them to recognize legitimate and illegitimate transactions with a high degree of accuracy. For instance, if a credit card transaction is made from an unusual location or for an atypically large amount, the AI system can flag it in real-time, potentially stopping fraud before it occurs.

In risk management, AI models predict potential loan defaults and assess credit risk by analyzing borrowers' spending behaviors, payment histories, and economic conditions. These insights help banks and lenders to make more informed lending decisions, optimize their portfolio exposures, and reduce the risk of bad debt.

Personalized Banking

AI drives personalization in banking through chatbots and virtual assistants, which provide customers with 24/7 assistance for a range of services such as account inquiries, financial advice, and transaction facilitation. These AI-driven tools use natural language processing to understand and respond to customer queries effectively, enhancing user experience and customer service.

Moreover, AI personalizes financial advice by analyzing individual customer data and offering tailored advice based on spending habits, income, and financial goals. For example, robo-advisors, which are automated platforms using AI algorithms, provide personalized investment management services with minimal human intervention, making financial advice more accessible and affordable.

Regulatory Compliance and Anti-Money Laundering

Compliance with financial regulations is critical, and AI assists institutions in adhering to these regulations more efficiently. AI systems can monitor and analyze transactions in real time to ensure compliance with regulatory requirements, reducing the risk of penalties and reputational damage. For example, in anti-money laundering (AML) efforts, AI can detect complex patterns and relationships across multiple accounts and transactions, identifying suspicious activities that would be difficult for human analysts to notice.

AI systems also help in reporting and auditing processes by automatically generating and aggregating necessary reports and documentation required by regulatory bodies. This automation not only saves time but also enhances the accuracy of reports and reduces human error.

3. Entertainment

AI finds fertile land in the entertainment industry, especially regarding:

Content Personalization and Recommendation Systems

AI has revolutionized the way content is delivered to users, particularly in streaming services such as Netflix, Spotify, and YouTube. These platforms use sophisticated AI algorithms to analyze vast amounts of data on user behavior, preferences, and viewing histories. The algorithms predict what kind of content a user may enjoy next, offering highly personalized content recommendations. This not only enhances user experience but also increases platform engagement and retention rates. The effectiveness of these recommendation engines is evident in the way users frequently discover new shows, movies, or music that align perfectly with their tastes.

Marketing and Advertising

AI significantly impacts marketing strategies within the entertainment industry. By analyzing data on consumer behavior and preferences, AI helps companies target their advertisements more effectively, ensuring that marketing campaigns reach the most receptive audiences. AI also enables dynamic ad insertion in streaming content, where ads are tailored to the viewer's profile, increasing the likelihood of engagement. Furthermore, sentiment analysis tools powered by AI examine social media and review sites to gauge public opinion on shows, movies, and games, providing producers and marketers with valuable feedback to guide future productions and promotions.

Content Creation and Curation

In content creation, AI tools are being used to generate music, video game environments, and even scripts. For instance, AI algorithms can analyze existing musical notes or screenplay styles and then use this information to create new compositions or dialogue in a similar vein. This capability is being explored in filmmaking and gaming to generate creative content quickly, assisting human creators by providing them with a base to refine and build upon. AI is also used in editing software to enhance audio and visual elements, automate color correction, and optimize editing workflows, significantly reducing the time and cost associated with post-production.

Ethical consideration on AI

Ethical considerations in artificial intelligence (AI) are essential for ensuring that its development and application benefit society while adhering to moral standards. According to M. Safdar, in “Ethical considerations in artificial intelligence”, significant ethical issue is the potential for AI systems to inherit or amplify biases present in their training data, leading to discriminatory outcomes in critical areas such as hiring, loan approvals, and law enforcement. Addressing this requires careful selection and ongoing evaluation of training data, along with mechanisms to identify and mitigate biases that emerge in AI behavior.

The complexity inherent in AI technologies, particularly those utilizing deep learning models, often results in systems that are not transparent, making it challenging for users and stakeholders to understand how decisions are derived. This opacity can be particularly problematic in areas like healthcare and criminal justice, where AI-driven decisions can significantly impact human lives. For instance, in healthcare, an AI system might determine patient treatment plans, and in criminal justice, it might influence parole or sentencing decisions. In such contexts, the inability to scrutinize or understand the decision-making process raises serious ethical concerns about fairness and accountability.

To address these issues, there is a growing emphasis on the development of explainable AI (XAI). XAI aims to create AI systems whose actions can be easily understood by humans. This involves designing AI models that not only make decisions but also provide insights into the reasoning behind these decisions. For example, rather than simply identifying a medical image as indicative of a disease, an explainable AI system would detail the specific features and patterns that led to its conclusion, such as highlighting areas of the image that exhibit potential signs of the condition.

Explainable AI is crucial for maintaining accountability as it allows users to verify the accuracy and fairness of the AI's decisions. In the judicial system, for example, an AI that assists in sentencing must transparently demonstrate how it has arrived at its recommendations to ensure that it is not perpetuating biases or errors that could lead to unjust outcomes.

Furthermore, the push for explainable AI also includes the development of standards and frameworks that guide how AI systems should articulate their decision-making processes. These guidelines are designed to ensure that explanations are provided in a manner that is meaningful and useful to the intended audience, whether they are technical experts, decision-makers in non-technical roles, or affected individuals..

Privacy concerns also arise with AI, especially with systems that process vast amounts of personal data. Ensuring the privacy of individuals involves implementing strong data protection measures and using techniques like differential privacy to safeguard personal information against unauthorized access and breaches.

Moreover, as AI and automation continue to evolve, they could displace workers in various sectors, raising concerns about job loss and economic inequality. This challenges policymakers to create solutions such as retraining programs, safety nets, and potentially new economic models like universal basic income to address the impacts of AI on employment.

Ensuring that AI respects human dignity and does not undermine human decision-making in personal or societal matters is also critical. As AI becomes more integrated into daily decision-making, it is important to maintain human oversight and ensure that AI supports rather than replaces human judgment.

Finally, the development and benefits of AI are not equally distributed globally, which could exacerbate existing disparities between nations and within societies. Promoting global cooperation

to ensure that AI technologies are developed and used in ways that are just and equitable is necessary to prevent widening these divides.

Addressing these ethical issues involves a multidisciplinary approach, drawing on expertise from fields such as law, philosophy, computer science, and public policy to develop AI technologies that are both innovative and responsible.

Challenges of interpreting and understanding this data (data transparency)

As AI technologies become increasingly sophisticated, they generate complex data that can be difficult to interpret and understand, presenting significant challenges, particularly regarding data transparency. Data transparency in AI refers to the clarity and understandability of data, processes, and decisions made by AI systems, which are essential for fostering trust and accountability. However, summarizing the ideas from Larsson and Heintz in “Transparency in Artificial Intelligence”, several obstacles impede transparency in AI applications:

Firstly, the inherent complexity of algorithms, especially in advanced models like deep learning, creates a 'black box' scenario where the decision-making processes are opaque even to their developers. This opacity can lead to mistrust or misinterpretation of the AI outputs among users. Furthermore, AI systems handle vast volumes of data that vary greatly—from structured numerical data to unstructured texts and images—making it difficult to track and analyze. Such complexity can obscure critical insights and complicate oversight and understanding.

Moreover, AI systems are susceptible to biases present in their training data. If unchecked, these biases can perpetuate and potentially amplify existing prejudices, leading to unfair outcomes and decisions. Ensuring transparency in how data is collected, processed, and used is crucial for identifying and mitigating such biases, thus upholding ethical standards.

“An explainable AI (XAI) system aims to make its behavior more understandable to humans by providing explanations. The terms interpretability and explainability are often used interchangeably, but they have different meanings. Explainability is more concerned with explaining the inner mechanism of the system (black-box model) or presenting the solution in

human-understandable terms. However, interpretability concerns finding causality between input and output or understanding the cause of a decision.”⁸

We are going to focus on the second term, interpretability, understanding how data visualization can enhance the interpretation of AI outputs.

⁸ Nagahisarchoghaei, M.; Nur, N.; Cummins, L.; Nur, N.; Karimi, M.M.; Nandanwar, S.; Bhattacharyya, S.; Rahimi, S. An Empirical Survey on Explainable AI Technologies: Recent Trends, Use-Cases, and Categories from Technical and Application Perspectives. *Electronics* 2023, 12, 1092.

Advantages of using data visualization to communicate AI-generated insights

As X. Wang asserts in his paper⁹, Data visualization plays a crucial role in enhancing the transparency of Artificial Intelligence systems, a critical aspect as these systems become increasingly integral to decision-making processes across various sectors. By transforming complex data sets and algorithmic processes into visually intuitive formats, data visualization makes AI systems more accessible and understandable to both technical and non-technical stakeholders. This accessibility is essential for demystifying the often opaque nature of AI operations, fostering greater trust and confidence in these technologies.

Visual representations of data sources, features, and distributions provide clear insights into the underlying structure and functionality of AI systems. By illustrating how data is organized, processed, and interpreted by AI, stakeholders can better understand the logic that drives AI decisions. This understanding is crucial for validating the AI's operations, ensuring that they align with expected ethical and operational standards.

Moreover, data visualization highlights the results, impacts, and broader implications of AI systems in a tangible form. It allows stakeholders to see the direct outcomes of AI processes, facilitating an evaluation of effectiveness, fairness, and accuracy. Interactive visualizations enable real-time engagement, where users can manipulate variables or conditions to see how changes in input affect AI outputs. This interactivity not only enhances user engagement but also empowers users to probe the robustness and reliability of AI systems.

⁹ Wang, X., Wu, Z., Huang, W. et al. VIS+AI: integrating visualization with artificial intelligence for efficient data analysis. *Front. Comput. Sci.* 17, 176709 (2023). <https://doi.org/10.1007/s11704-023-2691-y>

Crucially, visualization tools reveal gaps, imbalances, or outliers in the data that could potentially skew AI results. By identifying these issues, organizations can take corrective measures to refine AI models, thereby improving their performance and fairness. Data visualization also permits the comparison of AI-generated outcomes with alternative scenarios, benchmarks, or expectations, providing a comprehensive evaluation framework that assesses AI performance across different conditions.

In summary, data visualization is indispensable for achieving AI transparency. It bridges the gap between complex AI technologies and user comprehension, facilitating a deeper understanding and interaction with AI systems. Through effective visualization, AI becomes not just a tool for the few but a technology that is scrutinizable and influenceable by the many, enhancing its integration into society in a manner that is responsible, ethical, and inclusive.

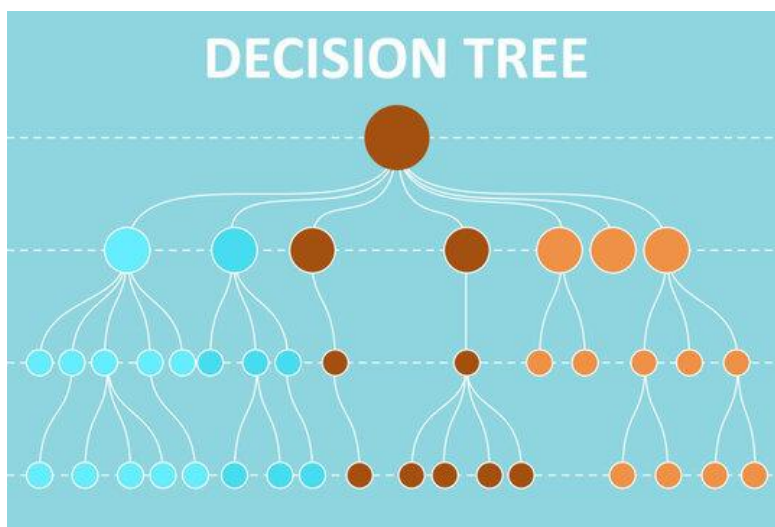
Better understanding of data through visualization tools

In the domain of artificial intelligence, where data is not only vast but also intricately complex, the role of visualization tools becomes fundamentally transformative. These tools enable a more profound comprehension of data by representing multidimensional datasets in forms that are visually intuitive and immediately actionable. By translating numerical and textual data into graphical formats—such as charts, graphs, and interactive simulations—visualization tools help elucidate underlying patterns and relationships that are not readily apparent from raw data alone. Moreover, visualization serves as a bridge between advanced analytical models and practical

business applications, allowing stakeholders to grasp sophisticated insights without the need for deep technical expertise. This accessibility is crucial for organizations aiming to leverage AI-driven analytics across diverse teams, fostering an environment where data-driven insights are at the core of strategic decision-making.

Additionally, visualization tools significantly enhance the interpretability of predictive models and algorithms, serving as a critical pillar of AI transparency. By transforming the outputs of intricate models from abstract numerical data into comprehensive visual narratives, these tools encapsulate the essence of the data, detailing its behaviors and the logic underlying its predictions. This visualization not only clarifies the results but contextualizes them within the framework of actionable insights.

For example, consider a complex machine learning model used to predict consumer behavior. A well-designed visualization, such as a series of decision trees, can vividly illustrate how different consumer attributes—like age, purchasing history, or geographical location—contribute to the model’s predictions. This type of visualization can unravel the decision-making process of the AI, showcasing the weight and impact of each input variable.



<https://cloud2data.com/what-is-decision-tree-in-machine-learning/>

Moreover, such graphical representations allow stakeholders to trace the decision path of the AI, from input through to outcome, enhancing their understanding of model accuracy and bias. By making the abstract tangible, visualizations facilitate a deeper engagement with the model's functionality, encouraging users to probe and question. This level of interaction is invaluable for refining AI systems, as it leads to more informed feedback loops where assumptions are tested and adjustments are made to improve model performance.

Furthermore, visual tools can highlight areas where the model may lack data or be overly influenced by certain features, prompting further investigation into potential biases or anomalies. This critical examination helps in building robust AI systems that are not only effective but also fair and reliable, thereby fostering a greater degree of trust among users and stakeholders.

In essence, by showing the complex interactions of algorithms and data, visualization tools do more than just report results—they create a story that highlights the detailed interactions of variables that drive AI systems. This story is crucial for making AI operations clearer, thereby improving transparency and strengthening trust in AI-driven decisions.

Thus, visualization tools do not merely simplify the visual presentation of data; they enhance the overall understanding, making AI insights more accessible, interpretable, and actionable. This capability is indispensable in harnessing the full potential of AI within any organization, driving both operational efficiencies and strategic innovations.

Facilitate decision making

Once again Wang using his model “VIS+AI” declares that data visualization not only clarifies complex datasets but also significantly enhances decision-making processes across various organizational levels. By presenting data in visual formats that are easy to understand and interpret, visualization tools help decision-makers quickly grasp key insights without needing to navigate through dense statistical data or complex algorithmic outcomes.

Visual tools like dashboards aggregate diverse data sources into a unified presentation, offering an at-a-glance view of critical metrics and performance indicators. This consolidated view facilitates a holistic understanding of situations, enabling executives, managers, and other stakeholders to make informed decisions rapidly. For instance, a financial dashboard might display real-time market trends, asset performance, and risk metrics simultaneously, allowing a financial analyst to make quick investment decisions based on comprehensive data insights.

Moreover, the use of interactive visualization tools allows for scenario analysis, where decision-makers can manipulate variables to see how different decisions might impact outcomes. This ability to model various outcomes based on real data underpins strategic planning and risk assessment, reducing the uncertainty inherent in decision-making processes. For example, a supply chain manager might use a visualization tool to simulate the impact of a change in supplier or a major increase in demand, helping to plan inventory levels and logistics more effectively.

In addition, visualization tools can democratize data access within an organization, empowering employees at all levels to engage with data directly and make decisions based on factual evidence. This empowerment leads to a more agile organizational culture, where faster and more decentralized decision-making can occur. For example, sales teams with access to real-time sales

data visualizations can adjust their strategies quickly based on what is most effective, improving both performance and responsiveness.

Furthermore, in environments where collaborative decision-making is crucial, visualization provides a common language that all stakeholders can understand, regardless of their technical expertise. This shared understanding fosters better collaboration and consensus-building as teams can discuss real-time data visualizations during meetings, ensuring that all voices are heard and that decisions are made with a comprehensive understanding of the underlying data.

In summary, data visualization is a critical tool that facilitates decision-making by providing clear, insightful, and interactive presentations of complex data. This capability not only speeds up the decision-making process but also enhances its quality by grounding decisions in solid data-driven evidence, promoting an informed, agile, and collaborative decision-making environment within organizations.

Case Studies

AI in Finance: role of Data Visualization

As we said before, “Machine learning (ML) is a subset of artificial intelligence that enables a system to autonomously learn and improve using neural networks and deep learning, without being explicitly programmed, by feeding it large amounts of data. It allows financial institutions to use the data to train models to solve specific problems with ML algorithms – and provide insights on how to improve them over time.”¹⁰

Summarizing what said previously, AI significantly impacts various aspects of the finance industry by enhancing efficiency, accuracy, and security. In algorithmic trading, AI uses machine learning algorithms to analyze large datasets rapidly, identifying patterns and executing trades with speed and efficiency unmatched by humans. This capability allows for adaptation to new information, optimizing returns and reducing risks in high-frequency trading environments.

AI also plays a crucial role in fraud detection and risk management. It analyzes transactional data to identify and prevent fraudulent activities, spotting anomalies that might indicate fraud. For instance, AI can flag potentially fraudulent credit card transactions in real-time based on unusual patterns such as location or transaction size. In risk management, AI models assess credit risk and predict loan defaults by analyzing data on borrowers' behaviors and economic conditions, aiding banks in making informed lending decisions and managing risks more effectively.

In personalized banking, AI enhances customer service through chatbots and virtual assistants, providing 24/7 assistance and using natural language processing to handle inquiries efficiently. AI

¹⁰ <https://cloud.google.com/discover/finance-ai>

also tailors financial advice by analyzing personal data, with robo-advisors offering customized investment strategies, thus making financial planning more accessible.

Furthermore, AI supports regulatory compliance and anti-money laundering efforts by monitoring transactions in real-time, ensuring adherence to laws and reducing the risk of penalties. It automates the generation and aggregation of compliance reports, enhancing the accuracy and efficiency of regulatory processes. This comprehensive integration of AI streamlines operations, improves security, and personalizes services in the finance sector.

But how Data visualization can help better understand this type of data?

For this section I have analyzed the article “Role of Data Visualization in Finance”¹¹ by Md. Kamrul Hasan Chy, Obed Nana Buadi, deep diving into the opportunities generated by Data Viz applied to Finance.

The article explores the transformative impact of data visualization on the comprehension of complex financial data. It asserts that data visualization serves as a critical tool for distilling vast and intricate financial information into digestible, visual formats that enhance clarity and facilitate more informed decision-making. Using practical examples, such as the financial disclosures of public companies like Walmart, the text demonstrates how charts, graphs, and dashboards can significantly aid stakeholders, including those with limited financial expertise, in understanding underlying financial trends and corporate health. It emphasizes that while traditional financial statements are laden with detailed numeric data and textual descriptions, incorporating visual aids can help highlight key financial metrics and trends more effectively, thus bridging the gap between complex financial data and stakeholder understanding. The section argues for a broader adoption of visual tools in financial reporting to not only streamline the analytical process but also to

¹¹ Chy, M. and Buadi, O. (2023) Role of Data Visualization in Finance. American Journal of Industrial and Business Management, 13, 841-856.

democratize financial information, making it accessible and understandable to a wider audience. This approach not only enhances transparency but also empowers stakeholders to make quicker and more accurate assessments of a company's financial standing.

Challenging and limitations of data visualization in the context of AI-generated data

Data visualization, while invaluable for interpreting AI-generated data, faces several challenges and limitations, particularly as the complexity and volume of AI data continue to escalate. One significant challenge is the visualization of high-dimensional data, which is common in AI applications. Traditional visualization techniques, such as scatter plots or histograms, struggle to effectively represent multidimensional relationships and patterns in a comprehensible manner. This can make it difficult for users to gain actionable insights from visualizations of complex machine learning models or deep learning operations.

Another limitation is the dynamic nature of AI data, which continuously evolves as AI systems learn and adapt from new data. Static visualizations often fail to capture these changes, potentially leading users to make decisions based on outdated information. While interactive and real-time visualizations attempt to address this, they require significantly more resources and sophisticated design to maintain accuracy and effectiveness.

Additionally, there is the challenge of scalability. As AI systems process increasingly vast datasets, visualizations must also scale without losing clarity or performance. This can be technologically demanding and may lead to oversimplified visual representations that omit critical nuances of the data.

Furthermore, ensuring that visualizations are not only informative but also interpretable to a diverse range of users, from AI experts to business stakeholders, adds another layer of complexity. This often requires customized visualization approaches that cater to different levels of technical expertise, which can be resource-intensive to develop.

Lastly, the risk of misinterpretation exists; visually represented data can be compelling yet misleading if not designed carefully. Misleading visualizations can arise from inappropriate use of scales, colors, or aggregation of data points, potentially leading to erroneous conclusions about the AI's performance or behavior.

Addressing these challenges requires ongoing innovation in visualization technologies and methodologies, along with a careful consideration of the intended audience and the specific characteristics of the AI-generated data being visualized.

Handling High-Dimensional Data

High-dimensional data refers to datasets that have a large number of variables or features for each observation, often much greater than the number of observations themselves. This type of data is common in many advanced fields such as machine learning, bioinformatics, and big data analytics, where each piece of data can contain hundreds or thousands of different measurements or attributes.

Characteristics of High-Dimensional Data

Many Features: The defining characteristic of high-dimensional data is that it includes a large number of variables or attributes. For example, in genomics, each sample might include measurements of the expression levels of thousands of different genes.

Sparsity: In many cases, high-dimensional data can be sparse, meaning that the majority of the measurements might be zero or near zero, indicating that most features do not have a significant impact or are not active in the observation.

Complexity in Visualization and Analysis: With so many dimensions, traditional visualization techniques and simple analytical methods become ineffective, making it difficult to discern patterns or insights directly from the data.

Curse of Dimensionality: As the number of dimensions increases, the amount of data needed to support meaningful analysis grows exponentially. This phenomenon, known as the "curse of dimensionality," can lead to models that overfit and perform poorly on new, unseen data unless properly managed.

Scalability Issues: challenges in scaling visualizations as data volume and real-time data streams increase

In the VIS+AI scenario proposed by Wang once again, we encounter some issues. First, scalability issues in data visualization pose significant challenges as the volume and velocity of data increase, particularly in the context of AI and big data applications. Performance degradation can occur as the computational load increases to process and render visualizations, leading to slow response times that are especially problematic in real-time scenarios like financial trading or network monitoring. Additionally, visual clutter can obscure important trends and insights when data points become overly dense, reducing the effectiveness of visual analysis. To manage large datasets while maintaining clarity, data aggregation is often employed, but this can result in a loss of important details and anomalies, diluting the granularity and potentially leading to oversight of crucial subtleties.

Maintaining responsive user interaction also becomes challenging as visualizations scale, with features such as zooming and filtering becoming slow, hindering effective data engagement and exploration. Furthermore, large-scale visualizations can exhaust significant memory and system resources, necessitating efficient resource management strategies to ensure sustainability and scalability. Ensuring accuracy and consistency in visual representations as data scales is critical, requiring rigorous data integrity and unbiased visual mappings.

To address these scalability challenges, several strategies are vital. Efficient data management through techniques like data indexing and powerful processing frameworks can enhance handling of large datasets. Employing advanced visualization techniques, such as progressive loading and level of detail rendering, helps manage large datasets dynamically. Utilizing cloud resources for computation and storage can also mitigate local system burdens, providing scalable solutions for

complex visualizations. Additionally, robust visualization libraries designed for big data can optimize performance and scalability, ensuring timely and accurate data insights. Addressing these scalability issues is crucial for leveraging large data volumes effectively, ensuring that visualizations remain a powerful tool for insight extraction in the digital age.

Future research

The future of data visualization in the context of AI-generated insights holds vast potential for innovation and improvement. One promising area for future research is the development of more sophisticated visualization techniques capable of handling the high dimensionality and complexity of AI data. This includes creating interactive and dynamic visualizations that can adapt in real-time to changing data streams, enhancing the user's ability to explore and understand AI outputs as they evolve.

Another critical area is the integration of explainable AI (XAI) methods with visualization tools to further demystify the 'black box' nature of advanced algorithms, thereby fostering greater trust and interpretability. Additionally, research should focus on improving the scalability of visualization tools to efficiently manage and present ever-growing datasets without sacrificing performance or clarity. The ethical implications of data visualization in AI, particularly concerning bias detection and mitigation, also warrant deeper investigation to ensure that visual representations are not only accurate but also fair and unbiased. Lastly, the development of user-centered design principles for AI visualization tools will be essential to cater to diverse audiences, from data scientists to business stakeholders, ensuring that these tools are intuitive and accessible to all. By addressing these areas, future research can significantly enhance the effectiveness of data visualization in making AI-generated insights more transparent, interpretable, and actionable.

Bibliography

- <https://www.linkedin.com/advice/3/how-can-data-visualization-help-ensure-rgijf#:~:text=By%20representing%20complex%20and%20large,are%20making%20predictions%20or%20decisions.>
- <https://learn.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview>, 22/04/2024
- Greg Allen, A concise, practical, and readable overview of Artificial Intelligence and Machine Learning technology designed for non-technical managers, officers, and executives, April 2020
- <https://www.cloudflare.com/learning/ai/what-is-predictive-ai/>, 23/07/2024
- C Janiesch, P Zschech, K Heinrich - Electronic Markets, 2021 - Springer, Machine learning and deep learning
- Feuerriegel, S., Hartmann, J., Janiesch, C. et al. Generative AI. Bus Inf Syst Eng 66, 111–126 (2024). <https://doi.org/10.1007/s12599-023-00834-7>
- Nagahisarchoghaei, M.; Nur, N.; Cummins, L.; Nur, N.; Karimi, M.M.; Nandanwar, S.; Bhattacharyya, S.; Rahimi, S. An Empirical Survey on Explainable AI Technologies: Recent Trends, Use-Cases, and Categories from Technical and Application Perspectives. Electronics 2023, 12, 1092. <https://doi.org/10.3390/electronics12051092>
- Mohammed Yousef Shaheen. Applications of Artificial Intelligence (AI) in healthcare: A review. ScienceOpen Preprints. 2021. DOI: 10.14293/S2199-1006.1.SOR-.PPVRY8K.v1
- Harikumar Pallathadka, Applications of artificial intelligence in business management, e-commerce and finance. Volume 80, Part 3, 2023, Pages 2610-2613
- <https://www.linkedin.com/pulse/role-artificial-intelligence-algorithmic-trading-theinvestorco-hfpnc/>. 25/04/2024

- <https://aicontentfy.com/en/blog/role-of-ai-in-content-curation>
- Longbing Cao, AI in Finance: Challenges, Techniques, and Opportunities, ACM Computing Surveys, Volume 55, Issue 3, 2022
- Chy, M. and Buadi, O. (2023) Role of Data Visualization in Finance. American Journal of Industrial and Business Management, 13, 841-856.
- Nabile M. Safdar, John D. Banja, Carolyn C. Meltzer, Ethical considerations in artificial intelligence, European Journal of Radiology, Volume 122, 2020
- Nassar, A., & Kamal, M. (2021). Ethical Dilemmas in AI-Powered Decision-Making: A Deep Dive into Big Data-Driven Ethical Considerations. International Journal of Responsible Artificial Intelligence, 11(8), 1–11.
- Wang, X., Wu, Z., Huang, W. et al. VIS+AI: integrating visualization with artificial intelligence for efficient data analysis. Front. Comput. Sci. 17, 176709 (2023). <https://doi.org/10.1007/s11704-023-2691-y>
- Muthukrishnan, Nikesh, et al. "Brief history of artificial intelligence." Neuroimaging Clinics of North America 30.4 (2020): 393-399.