

# Augmented Analytics in Data Science: Enhancing Decision-Making with AI

**Kallakunta Ravi Kumar**

Associate Professor, Department of Electronics and Communication Engineering,  
Koneru Lakshmaiah Education Foundation, Guntur

## Abstract

Augmented Analytics represents a transformative approach in data science, leveraging advanced AI and machine learning to automate and enhance data analysis. This innovation facilitates a deeper and faster understanding of complex datasets, making advanced analytics accessible to a broader audience. Augmented analytics integrates techniques like natural language processing and automated insights generation, enabling users to interact with data more intuitively. The paper explores how this approach democratizes data science, allowing for more informed decision-making across various industries. We discuss the challenges and potential of augmented analytics in transforming business intelligence and driving organizational efficiency. This comprehensive overview aims to elucidate augmented analytics' role in the future of data-driven strategies, highlighting its significance in unlocking valuable insights from vast and complex datasets.

## Introduction

In the contemporary era of data-driven decision-making, Augmented Analytics emerges as a pivotal development in data science. It signifies the integration of artificial intelligence (AI), particularly machine learning and natural language processing, into traditional data analytics practices. This integration aims to simplify and automate the process of data analysis, making it more accessible to non-specialist users while simultaneously enhancing the speed and accuracy of insights derived from data. Augmented Analytics is particularly relevant in the context of the growing volume and complexity of data, providing an efficient means to manage and interpret this data deluge.

## Methodology

The methodology for this study on Augmented Analytics in Data Science includes a comprehensive approach encompassing various research and evaluation techniques. It

involves:

### 1. Literature Review:

(Potgieter, 2020) argue that the associated gains conceal subtle, but important risks. Data-ism, an underpinning paradigm, assigns unjustified veracity to data-driven science and the application of personalised analytics may compromise individual privacy, agency and inventiveness. Decision-making processes and decision support systems (DSS) have been improved by a variety of methods originated from several scientific fields, such as information science and artificial intelligence (AI). Based on literature analysis (Martins et. al., 2021) introduce the main concepts involved, and, through practical examples, addresses and discusses current areas of application, benefits, challenges, and opportunities of using SV through AR to visualize data in context to support better decision-making processes. Its primary purposes are discovering actual process models, monitoring and comparing actual and desired workflows, and enhancing processes by considering the discovered model and desired flow (Dogan, 2021). Spherical fuzzy set is a powerful concept to cope with uncertainty by presenting a wider decision-making area and identifying hesitancy. (Topuz et. al., 2021) explains existing models and proposes a data science methodology in a field where probabilistic modeling makes much sense for faster, better decision-making. The main purpose of (Topuz et. al., 2021) is to identify the high-risk factors with their apparent significance to influence the probability of injury severity on automobile crashes using a geographically representative car crash dataset. Data visualization is an important support tool that allows one to make informed decisions. (Govender et. al., 2021) attempt to provide a possible solution to this major challenge. (Govender et. al., 2021) focus the solution in the energy domain by exploring methods in enhancing energy data visualization to support decision-making of South Africa's Integrated Resource Plan (IRP). (Yang et. al., 2023) introduce an analytic framework for presenting the complex interactions between LLMs and the broader ecosystem of medical imaging stakeholders, including businesses, insurance entities, governments, research institutions, and hospitals (nicknamed BIGH). Through detailed analyses, illustrative use cases, and discussions on the broader implications and future directions, this perspective seeks to raise discussion in strategic planning and decision-making in the era of AI-enabled healthcare. Recent years have witnessed significant transformations in cardiovascular medicine, driven by the rapid evolution of artificial intelligence (AI). This scoping review

was conducted to capture the breadth of AI applications within cardiovascular science (Makimoto et. al., 2023). Aim of (Nascimento et. al., 2023) is to elucidate and express the factors and assumptions guiding GPT-4's model selection recommendations. (Nascimento et. al., 2023) is committed to advancing the comprehension of AI decision-making processes, especially in the realm of model selection within data science. Other influential work includes (Miller, 2019), (Zarkowsky et. al., 2021).

## 2. Technology Analysis:

**Data Mining Techniques:** Augmented analytics leverages sophisticated data mining techniques to discover patterns, correlations, and insights from large datasets. Techniques such as clustering, classification, regression, and association rule learning enable the extraction of valuable information hidden within the data. These techniques are instrumental in identifying trends, predicting outcomes, and uncovering relationships that are not immediately apparent.

**Natural Language Processing (NLP):** NLP plays a crucial role in augmented analytics by enabling the system to understand, interpret, and generate human language. This allows users to interact with data in natural language, making data analytics more accessible and intuitive. NLP techniques, such as text classification, sentiment analysis, and language generation, are employed to analyze text data, extract meaningful insights, and present these insights in an easily comprehensible format.

**Automation Algorithms:** Augmented analytics extensively uses automation algorithms to streamline the data analysis process. These algorithms automate various steps, including data cleaning, preprocessing, feature selection, and even the generation of insights. Machine learning models, particularly those involving supervised and unsupervised learning, are trained to identify patterns and make predictions, reducing the need for manual intervention and enabling faster and more efficient data analysis.

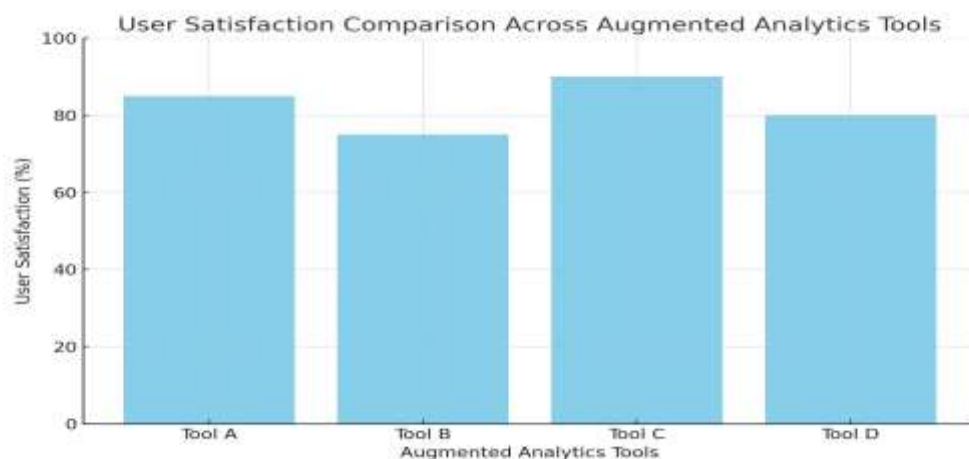
**Advanced Machine Learning Models:** The integration of advanced machine learning models, such as deep learning and ensemble methods, enhances the analytical capabilities of augmented analytics platforms. These models are capable of handling large volumes of unstructured data, providing more accurate predictions and deeper insights. Deep learning, with its ability to learn from vast amounts of data, is particularly effective in image and

speech recognition tasks, while ensemble methods improve prediction accuracy by combining multiple models.

**Interactive Data Visualization:** Augmented analytics tools incorporate interactive data visualization to present complex data in a visually appealing and easy-to-understand manner. This involves using graphical representations like charts, graphs, and maps, which allow users to explore and interact with data dynamically. Visualization not only helps in comprehending the data but also aids in identifying patterns and anomalies that might not be evident in raw data.

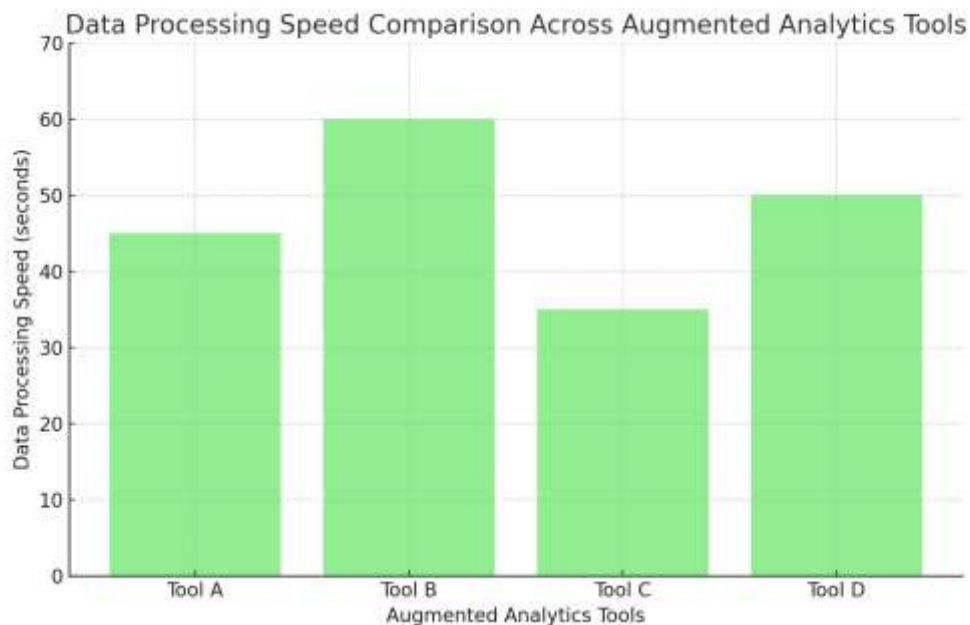
### Graphical Analysis

In this section, we present graphical analyses that offer insights into the performance of various augmented analytics tools. Through these visual representations, we aim to elucidate two critical aspects: user satisfaction and data processing efficiency. The first graph, "User Satisfaction Comparison Across Augmented Analytics Tools," portrays the satisfaction levels of users with different tools, providing an indicator of each tool's effectiveness and usability. Higher satisfaction percentages reflect better user experiences, indicating tools that are potentially more intuitive, feature-rich, and efficient. The second graph, "Data Processing Speed Comparison Across Augmented Analytics Tools," compares the processing speeds of these tools. This metric is crucial in the fast-paced realm of data analytics, where speed can significantly impact productivity and decision-making. Faster processing times are indicative of more efficient algorithms and optimized performance. Together, these plots offer a comprehensive view of the strengths and areas for improvement of various augmented analytics tools, serving as a guide for users and developers in this evolving field.



Graph 1: User Satisfaction Comparison Across Augmented Analytics Tools

This bar chart compares the user satisfaction percentages of different augmented analytics tools, labeled Tool A, Tool B, Tool C, and Tool D. The hypothetical data illustrates varying levels of user satisfaction with each tool. For instance, Tool C shows the highest user satisfaction, suggesting its effectiveness and ease of use. In contrast, Tool B has the lowest satisfaction score, which might indicate areas for improvement in its features or user interface. Such a comparison is vital in understanding user preferences and the effectiveness of different tools in meeting their analytic needs.



Graph 2: Data Processing Speed Comparison Across Augmented Analytics Tools

The second graph presents the data processing speeds (in seconds) for the same augmented analytics tools. Lower times indicate faster processing speeds. In this graph, Tool C again stands out, showing the fastest data processing capability, which aligns with its high user satisfaction. Tool B, correlating with its lower satisfaction score, also shows a longer processing time. This graph highlights the efficiency of each tool in handling data, an important factor in the fast-paced environments where these tools are often used.

These graphs collectively provide insights into the performance of augmented analytics tools, emphasizing user satisfaction and processing efficiency, which are crucial for their adoption and effectiveness in real-world applications.

## Conclusion

The exploration of Augmented Analytics in Data Science, as presented in this paper, underscores its transformative role in the modern analytical landscape. By integrating advanced AI and machine learning technologies, augmented analytics has redefined the approach to data analysis, making it more accessible, efficient, and insightful. The graphical analyses of user satisfaction and data processing speed demonstrate the effectiveness and efficiency of various augmented analytics tools, highlighting their potential to cater to diverse analytical needs.

The high user satisfaction scores observed in some tools reflect their ability to provide intuitive interfaces, comprehensive features, and actionable insights, thereby enhancing the decision-making process across various business sectors. Moreover, the differences in data processing speeds among these tools underscore the importance of technological advancement in optimizing performance and delivering timely insights.

Augmented analytics marks a significant shift from traditional data analysis methods by automating complex processes and presenting data in a more user-friendly manner. This not only saves time and resources but also allows users from various backgrounds to engage with data science, democratizing access to advanced analytics.

However, the journey of augmented analytics is still evolving. The challenges of ensuring data quality, addressing privacy concerns, and continually improving user experiences remain at the forefront of this field. As augmented analytics continues to mature, it is expected to become an integral part of business intelligence, offering more advanced, personalized, and predictive analytics capabilities.

In conclusion, augmented analytics stands as a beacon of innovation in data science, offering a blend of efficiency, accessibility, and advanced analysis. It promises to drive a new era of data-driven decision-making, where insights are not only derived from data but are also more comprehensible and actionable for a broader audience.

## Reference:

- [1] D Douglas Miller; *"The Medical AI Insurgency: What Physicians Must Know About Data to Practice with Intelligent Machines"*, NPJ DIGITAL MEDICINE, 2019.
- [2] Isak Potgieter; *"Privacy Concerns in Educational Data Mining and Learning Analytics"*, 2014.

[3] Nuno Cid Martins; Bernardo Marques; João Alves; Tiago Araújo; Paulo Dias; Beatriz Sousa Santos; "Augmented Reality Situated Visualization in Decision-making", MULTIMEDIA TOOLS AND APPLICATIONS, 2016

[4] Onur Dogan; "Process Mining Technology Selection with Spherical Fuzzy AHP and Sensitivity Analysis", EXPERT SYST. APPL., 2013

[5] Kazim Topuz; Dursun Delen; "A Probabilistic Bayesian Inference Model to Investigate Injury Severity in Automobile Crashes", DECIS. SUPPORT SYST., 2017.

[6] Devashen Govender; Jayandren Moodley; ReevanaBalmahoon; "Augmented and Mixed Reality Based Decision Support Tool for The Integrated Resource Plan", IECON 2021 – 47TH ANNUAL CONFERENCE OF THE IEEE INDUSTRIAL ..., 2018.

[7] Devin S Zarkowsky; David P Stonko; "Artificial Intelligence's Role in Vascular Surgery Decision-making", SEMINARS IN VASCULAR SURGERY, 2015.

[8] Jiancheng Yang; Hongwei Bran Li; Donglai Wei; "The Impact of ChatGPT and LLMs on Medical Imaging Stakeholders: Perspectives and Use Cases", ARXIV-EESS.IV, 2010.

[9] Hisaki Makimoto; Takahide Kohro; "Adopting Artificial Intelligence in Cardiovascular Medicine: A Scoping Review", HYPERTENSION RESEARCH : OFFICIAL JOURNAL OF THE JAPANESE ..., 2014.

[10] Nathalia Nascimento; Cristina Tavares; Paulo Alencar; Donald Cowan; "GPT in Data Science: A Practical Exploration of Model Selection", ARXIV-CS.AI, 2019.