iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80847

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Dual Minds: AI Agents Collaborating for Data-Driven Solutions

Gayatri Sharad Thombare[1], Dr. Rashmi Thakur[2]
Department of Computer Engineering, Thakur College of Engineering & Technology, Maharashtra, India

Abstract: Organisations in a variety of industries, including retail and healthcare, have depended more and more on data-driven methods in recent years to aid in both strategic and operational decision-making. While traditional data analytics and forecasting methods provide useful information, they often require expert interpretation and cannot provide non-technical users with useful recommendations. Furthermore, existing analytical systems are typically domain-specific and unadaptable when used with heterogeneous datasets. This work describes a multi-domain decision intelligence system that combines forecasting, AI-based advice, and automated exploratory data analysis (EDA) using large language models (LLMs) and retrieval-augmented generation (RAG). The proposed system automatically identifies an incoming dataset's domain, infers its schema, and generates understandable data summaries without requiring human configuration. Machine learning-based forecasting is used to identify non-linear trends in domains with rich historical patterns, such retail sales; solid baseline forecasting techniques are utilised for other domains, like healthcare, to guarantee dependability and interpretability .The system integrates domain-specific knowledge obtained from RAG with deterministic EDA insights and forecast summaries to close the gap between analytics and decision-making. An LLM is then given this grounded contextual data to produce trustworthy, comprehensible, and context-aware natural language suggestions .The suggested paradigm minimises hallucinated reactions, facilitates informed decision-making, and adapts well across domains, according to experimental evaluation on retail and healthcare datasets. The findings demonstrate the potential of combining RAG-enabled LLMs with structured data analytics to create reliable and user-focused decision support systems.

Keywords: Decision Intelligence, Exploratory Data Analysis, Forecasting, Machine Learning, Large Language Models, Retrieval-Augmented Generation, Multi-Domain Analytics

I. INTRODUCTION

Data-driven decision-making has grown in importance in recent years in a variety of industries, such as retail, healthcare, finance, and logistics. Organisations collect enormous volumes of structured data using sensors, digital platforms, and transactional systems. This data could provide useful insights into operational performance, resource usage, and client behaviour. However, it is still very challenging to transform raw data into actionable decisions, particularly for stakeholders who are not technical. Traditional data analytics solutions are primarily focused on descriptive reporting through dashboards and visualisations. These systems are helpful for summarising historical trends, but they need domain expertise to assess results and do not inherently support predictive or prescriptive decision-making. Consequently, there is a gap between data analysis and practical decision-making. In order to forecast future patterns using past data, forecasting techniques have been thoroughly researched. In time-series and demand forecasting challenges, machine learning models—particularly ensemble techniques like Random Forests—have proven to be highly effective in capturing non-linear correlations. These models are frequently used in the retail industry to forecast seasonal variations, inventory demand, and sales volume. Forecasting models are not enough to make decisions on their own, even though they can make predictions. In order to obtain useful insights, forecast results are usually numerical and need to be interpreted by specialists. Furthermore, it is challenging to generalise forecasting methods across heterogeneous data sources because most of them are created for particular domains and datasets.

The restrictions are much more noticeable in fields like healthcare. Complex machine learning models may be less reliable in healthcare datasets due to their high unpredictability, irregular patterns, and shallow historical depth. As a result, the decision-support capabilities of many healthcare analytics systems are further limited by their reliance on descriptive analysis or basic statistical techniques. New opportunities for natural language engagement with data have been made possible by recent developments in large language models (LLMs). LLM-based systems reduce the barrier to data access by allowing users to query datasets and analytical results through conversational interfaces. These systems have been investigated for decision support, report production, and conversational analytics.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

LLMs do not have direct access to domain-specific data and organisational expertise because they are trained on extensive general-purpose corpora. Because of this, LLMs may produce responses that are believable but factually false—a phenomenon known as hallucination[8].

In decision-critical applications, where inaccurate suggestions may result in unfavourable results, this constraint presents a serious danger. Retrieval-Augmented Generation (RAG) has been proposed as a mechanism to anchor language model outputs in external knowledge sources in order to address the dependability issues of LLMs. By obtaining pertinent documents or domain-specific knowledge and integrating it into the model's context during inference, RAG improves the generation process. In knowledge-intensive activities, this method has been demonstrated to improve factual consistency and lessen hallucinations.

Although RAG has proven successful in document-centric and question-answering applications, its integration with structured data analytics and forecasting is yet largely unexplored. Rarely do current systems integrate RAG-based reasoning, predictive modelling, and automated data analysis into a single decision intelligence framework[15].

The observation that existing analytical methods only handle portions of the decision-making pipeline serves as the driving force for this study. While forecasting models concentrate on prediction accuracy, dashboards concentrate on visualisation, and LLM-based systems concentrate on interaction, none of these elements by themselves offer comprehensive decision support. Furthermore, when applied to different datasets, the majority of current solutions need manual adaption because they are strongly connected to a single domain .Adaptive systems that can automatically comprehend dataset properties, produce comprehensible insights, predict future trends, and convert analytical results into practical suggestions are becoming more and more necessary. Non-technical users should be able to utilise these systems without sacrificing their dependability and explainability.

This study suggests a multi-domain decision intelligence system that combines forecasting, RAG-based large language models, and exploratory data analysis to address these issues. The following are this work's main contributions:

1) Multi-Domain Adaptability: Without the need for manual configuration, the suggested system may operate seamlessly across diverse datasets by automatically identifying dataset domains and inferring schema information.
2) Integrated Analytics and Forecasting: The system integrates domain-specific forecasting techniques with automated EDA, utilising machine learning for data-rich domains and reliable baseline techniques for data-poor domains.
3) RAG-Grounded Advisory Framework: The system generates dependable and comprehensible recommendations while reducing hallucinogenic responses by using domain-specific knowledge through Retrieval-Augmented Generation.
4) User-Centric Decision Support: The framework bridges the gap between data analytics and natural language explanations and suggestions and practical decision-making for non-technical users.

II. LITERATURE REVIEW

Data-driven decision-making has been extensively researched in a variety of fields, including operations management, healthcare, and retail. It is widely acknowledged that exploratory data analysis (EDA) is an essential step in comprehending the structure of datasets, spotting trends, spotting anomalies, and directing further modelling choices. EDA improves analytical reliability by giving analysts intuition about data distributions and relationships before predictive modelling, according to both classical and applied research [1]. However, traditional EDA methods are less helpful for non-technical decision-makers and real-time decision support because they mainly rely on static visualisations and human interpretation. Machine learning approaches have been widely used to solve forecasting issues, especially in retail demand prediction. When it comes to identifying non-linear patterns, seasonal variations, and feature interactions in sales data, ensemble-based models like Random Forests have proven to be highly effective [2]. Predictive models can be integrated into analytical platforms to facilitate planning and optimisation tasks, as demonstrated by recent applied machine learning systems [3]. Despite these developments, the majority of research on forecasting focuses on predictive accuracy rather than how forecast outputs can be used to make practical decisions. Because of this, decision-makers frequently need professional assistance in order to understand numerical forecasts. The difficulties associated with healthcare analytics are distinct. According to current research, healthcare datasets frequently display high variability, limited historical depth, and irregular temporal behaviour, all of which lower the efficacy and dependability of sophisticated machine learning models [4]. In order to preserve interpretability and operational safety, many healthcare analytics systems prioritise descriptive analysis or conservative baseline forecasting techniques. Although these approaches offer valuable insights, they are still isolated from higher-level decision-making processes. Large language models (LLMs), which allow users to communicate with analytical systems through natural language queries, have recently been investigated for conversational analytics and decision support [5]. These systems increase user engagement and accessibility, but numerous studies have revealed that LLMs are prone to hallucinations, producing responses that seem plausible but are not supported by actual data [6].

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

The application of LLMs in data-critical decision environments is severely limited by this constraint. The potential of language models as reasoning agents capable of supporting complex decision-making tasks is highlighted by recent studies on agent-based systems and multi-agent collaboration using LLMs [8]. Although these methods investigate sophisticated reasoning and coordination skills, their applicability in operational decision systems is diminished because they frequently lack close integration with deterministic data analytics and predictive modelling components. From an engineering standpoint, contemporary AI system design places a strong emphasis on separating probabilistic elements like language-based reasoning and explanation from deterministic elements like data pre-processing, analytics, and forecasting [9]. In practical AI applications, this architectural principle promotes dependability, interpretability, and maintainability. Few current systems, nevertheless, completely embrace this paradigm in the context of multi-domain decision intelligence. Exploratory analysis, forecasting, conversational engagement, and knowledge grounding are generally addressed in isolation by current methodologies, according to the literature. Unified frameworks that automatically adjust to various domains, incorporate forecasting and EDA outputs, and use RAG-enabled LLMs to provide dependable, comprehensible, and useful decision suggestions are lacking. In order to overcome these constraints, this study suggests a multi-domain decision intelligence system that integrates RAG-grounded advice, automated EDA, and adaptive forecasting techniques into a single, coherent framework.

III. SYSTEM ARCHITECTURE AND METHODOLOGY

The proposed system is a multi-stage, modular framework for decision intelligence that combines forecasting, data analytics, and AI-driven advice into a single adaptive pipeline. By emphasizing the separation of concerns, the architecture makes sure that probabilistic language-based reasoning is clearly separated from deterministic data processing and predictive components. Across a variety of application domains, this design enhances extensibility, explain ability, and reliability. Data ingestion and pre-processing, domain and schema inference, exploratory data analysis and forecasting, insight and context generation, and AI-based decision advisory comprise the system's five sequential stages. Every step gradually helps turn unprocessed datasets into recommendations that can be put into practice.

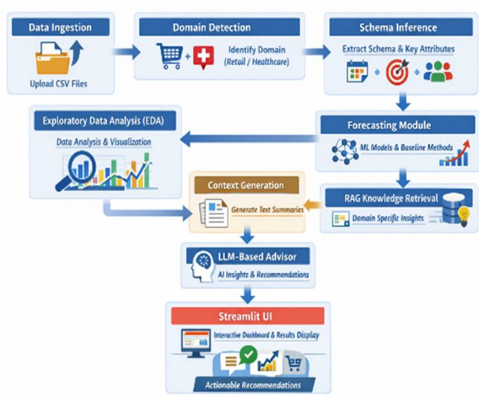


Figure 1 System architecture

Pre-processing and Data Ingestion : Structured datasets in comma-separated value (CSV) format are first accepted by the system. Column names are normalised during ingestion to guarantee uniformity across diverse datasets. In order to facilitate further analytical operations, date fields are parsed into standard temporal formats and numerical target variables are verified. The system can reliably handle datasets from various domains thanks to this pre-processing step, which eliminates the need for manual reconfiguration.

Detecting the Domain and Inferring the Schema: The system automatically figures out what domain the input dataset belongs to, like retail or healthcare, based on the semantics of the columns and the characteristics of the data. This makes it easier for the system to work in multiple domains. After the domain is found, a schema inference module finds the most important structural parts of the dataset, such as the temporal column, target variable, and relevant grouping attributes (like product identifiers in retail or department identifiers in healthcare).This automated schema inference allows the analytical and forecasting modules to operate generically across datasets, eliminating the need for hard-coded assumptions. By dynamically adapting to dataset structure, the system maintains flexibility and scalability across domains.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Exploratory Data Analysis and Forecasting :Once the dataset structure is identified, the system performs automated exploratory data analysis (EDA). This includes computing summary statistics, identifying dominant groups, and analyzing temporal trends. Visual analytics such as line charts, bar charts, and rolling averages are generated to provide an intuitive understanding of historical behavior.

For forecasting, the system applies domain-appropriate strategies. In domains with rich historical patterns, such as retail sales, a machine learning–based forecasting approach is employed. An ensemble regression model is trained on historical data to capture non-linear temporal patterns and demand fluctuations. In contrast, for domains such as healthcare, where data may be limited or irregular, the system applies robust baseline forecasting techniques to ensure interpretability and avoid unreliable predictions. The forecasting outputs are stored and reused across system components, enabling both visualization and downstream reasoning.

Insight and Context Generation :To bridge the gap between numerical analytics and decision reasoning, the system generates deterministic textual insights based on EDA and forecasting results. These insights describe observed trends, dominant contributors, and projected behaviors in natural language. This stage ensures that key analytical findings are explicitly captured as factual context rather than inferred implicitly by the language model. In parallel, forecast summaries are computed to provide concise representations of future projections. The combination of analytical insights and forecast summaries forms a structured context that serves as a reliable foundation for decision advisory.

RAG-Based Advisory Decisions : The system's last phase incorporates a big language model based on Retrieval-Augmented Generation (RAG) to offer suggestions for decisions. Domain-specific knowledge documents are indexed and retrieved using similarity-based search mechanisms. Relevant knowledge is injected into the model’s prompt alongside the structured analytical context generated in earlier stages. The language model's reasoning is grounded in both external domain knowledge and data-driven insights, which reduces hallucinated responses and increases suggestion dependability. By converting structured context into practical advice delivered in normal language, the language model serves as a layer of reasoning and explanation.

User Interface and Implementation :A modular Python-based architecture is used to create the system, which is then made available via an interactive web interface. Non-technical users can upload datasets, examine analytical findings, see forecasts, and communicate with the AI adviser using natural language enquiries thanks to the user interface. Future improvements and further domain integration are supported by the modular design, which permits the extension or replacement of specific components without impacting the system as a whole.

IV. EXPERIMENTAL SETUP AND RESULTS

A. Datasets Description

Datasets from two distinct application domains—retail sales and healthcare operations—were used to assess the suggested system. These datasets were chosen to show how the framework can be applied to a variety of data types .Dates, product identifiers, store details, and units sold are among the historical transactional records included in the retail dataset. The dataset is appropriate for machine learning-based forecasting because it displays temporal patterns like seasonality and demand fluctuations. The operational data pertaining to patient flow and departmental workload is represented by the healthcare dataset. Healthcare data, in contrast to retail data, exhibits erratic temporal behaviour and shallow historical depth, which encourages the application of conservative baseline forecasting techniques to guarantee interpretability and dependability.

Datasets	Domain	Key attribute
Retail sales	sales	Date ,Product, Units sold
Healthcare	Healthcare	Date ,Department, Patient count

Table 1 :summarize datasets characteristics

B. Experimental Configuration

Every experiment was carried out with an implementation based on Python. Data pre-processing, forecasting, exploratory data analysis, and decision advisory system components were carried out in the order specified by the system architecture. A forecasting method based on machine learning was used for the retail industry. An ensemble regression model that can capture non-linear demand patterns was trained using historical sales data. Line and bar charts were used to display the forecasts, which were created for a user-specified future horizon. A baseline forecasting approach was used for the healthcare sector. This method prioritises interpretability over predictive complexity by projecting future values based on aggregated statistics and recent historical trends. The operational limitations and risk sensitivity of healthcare data are in line with this design decision.

Data-driven insights, forecast summaries, and domain-specific knowledge obtained via Retrieval-Augmented Generation were combined to assess the decision advisory module. This structured context served as the foundation for the recommendations produced by the large language model.

C. Findings from Exploratory Data Analysis

For both datasets, automated exploratory data analysis was carried out. EDA demonstrated distinct temporal trends and product-level contribution patterns in the retail industry. Dominant products and seasonal behaviour were highlighted by visualisations like daily sales trends and aggregated product distributions. EDA results highlighted temporal variations in patient load and departmental variability in the healthcare sector. Resource planning decisions were supported by group-wise aggregation, which revealed departments with a consistently higher workload.

Figure 2 forecast plots for retail sales

D. Predictive Outcomes

The forecasting module generated demand projections for the retail industry that were consistent with past patterns. The system was able to support supply chain and inventory planning decisions by estimating expected sales volume for future periods using the forecasts that were generated. Based on recent historical data, baseline forecasts offered reliable patient load projections for the healthcare sector. These forecasts provide understandable and secure estimates appropriate for high-stakes operational environments, despite their inability to capture complex dynamics.

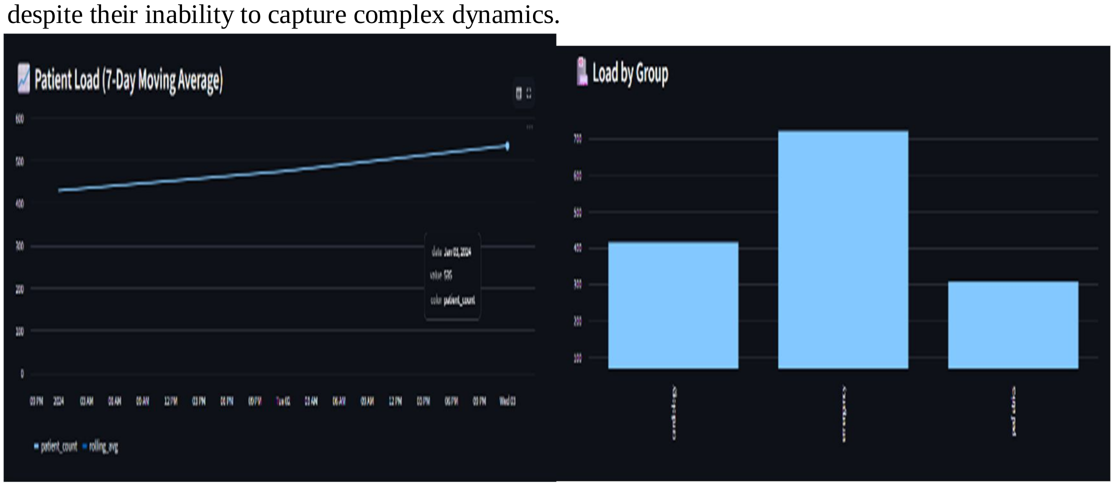


Figure 3 Baseline healthcare forecasts.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

E. Results of Decision Advisory

Using forecast summaries, retrieved domain knowledge, and EDA insights, the decision advisory component produced natural language recommendations. The advisor offered advice on demand patterns and inventory adjustment in the retail industry. Recommendations in the healthcare sector centred on staffing issues and workload distribution. Response reliability was considerably increased by grounding the language model using Retrieval-Augmented Generation, according to qualitative evaluation. The RAG-enabled advisor showed better alignment with observed data patterns and less hallucination than LLM-only responses.

Figure 4 Advisor responses illustrating grounded recommendations.

F. Results Discussion

The experimental findings show that the suggested framework successfully combines forecasting, analytics, and AI-based advisory across various domains. The system maintains a consistent decision advisory interface while adjusting its forecasting and analytical strategies according to dataset characteristics .RAG-grounded language modelling in conjunction with deterministic data analysis guarantees that recommendations continue to be dependable and explicable. These findings confirm that implementing multi-domain decision intelligence systems that connect data analysis and practical decision-making is feasible.

V. CONCLUSION AND FUTURE WORK

In order to facilitate data-driven decision-making, this study introduced a multi-domain decision intelligence system that combines forecasting, big language models based on Retrieval-Augmented Generation, and exploratory data analysis. The suggested methodology combines deterministic analytics with grounded language-based reasoning to bridge the gap between data analysis and actionable decision assistance, in contrast to conventional analytical dashboards or stand-alone forecasting models. By identifying domain features and deriving schema information, the system automatically adjusts to diverse datasets, facilitating smooth operation across many domains including retail and healthcare. The system successfully creates interpretable insights, relevant forecasts, and trustworthy natural language recommendations based on both domain-specific knowledge and data-driven summaries, according to experimental evaluation. Retrieval-Augmented Generation greatly decreased hallucinogenic responses and increased the advisory outputs' credibility.

The suggested design guarantees explainability and dependability, which are essential for real-world decision-making situations, by isolating probabilistic language production from analytical processing. The outcomes demonstrate how feasible it is to combine RAG-enabled large language models, automated analytics, and machine learning into a single, user-focused decision support system.

A. Future Work

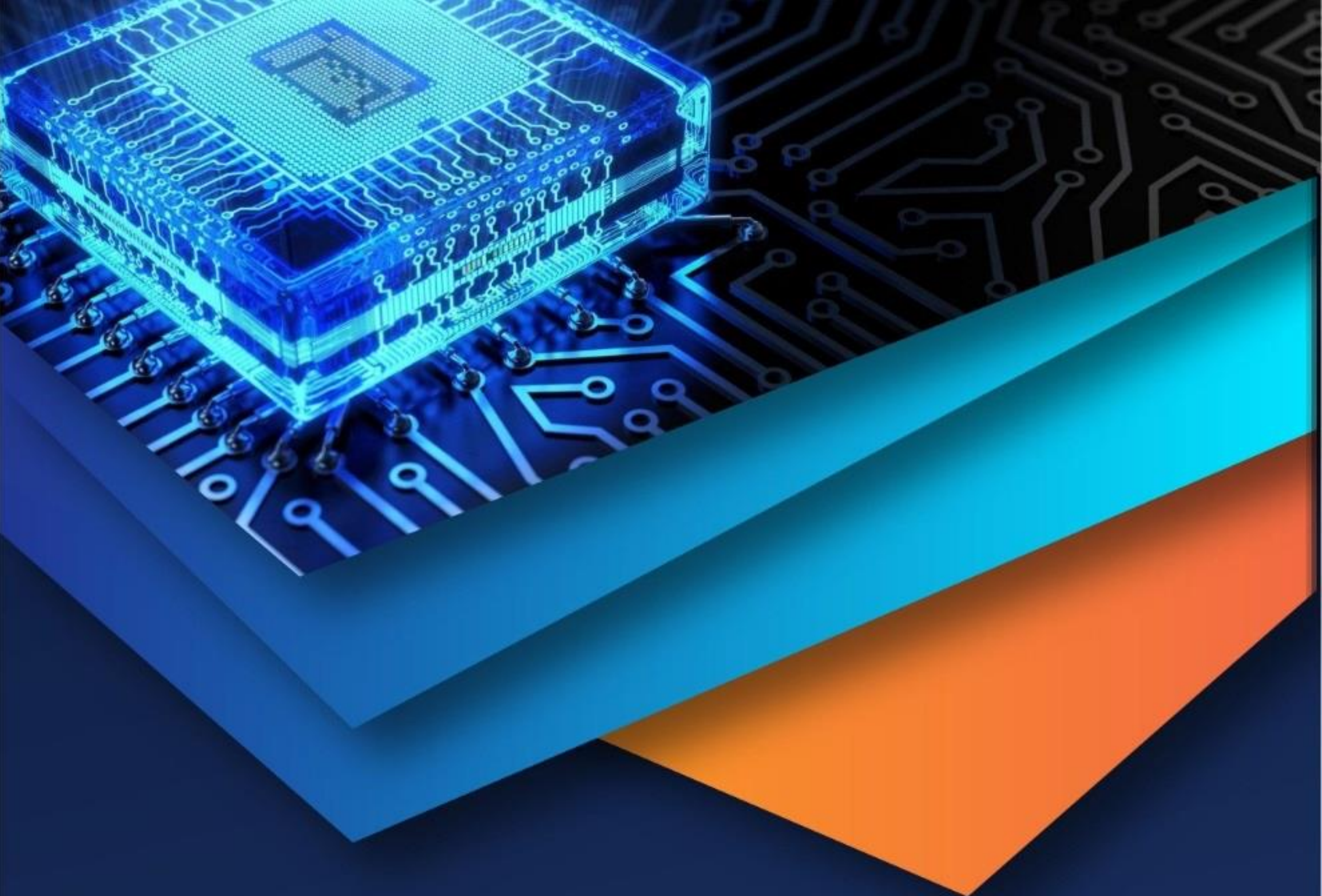
Although the suggested framework shows encouraging outcomes, there are a few ways to improve it. When enough domain-specific data is available, further studies might use deep learning-based techniques or sophisticated time-series forecasting models like Prophet. To further assess system generalisability, other application domains such as banking, education, and logistics might be included.

Furthermore, to measure prediction reliability, future studies can investigate confidence estimation and uncertainty-aware forecasting. Another possible approach is to improve the decision advising component using adaptive knowledge updating techniques and user feedback loops. The system's function as a reliable and scalable platform for decision intelligence can be strengthened by these additions.

REFERENCES

[1] J. W. Tukey, "Exploratory Data Analysis," Addison-Wesley, 1977.
[2] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[3] ECML-PKDD, "Applied Machine Learning Systems for Data Analytics," European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 2025.

[4] P. Harper, “Healthcare Analytics: Methods and Applications,” Springer, 2018.
[5] T. Brown et al., “Language Models are Few-Shot Learners,” Advances in Neural Information Processing Systems (NeurIPS), 2020.
[6] Z. Ji et al., “Survey of Hallucination in Natural Language Generation,”Association for Computational Linguistics (ACL), 2023.
[7] P. Lewis et al., “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks,” Advances in Neural Information Processing Systems (NeurIPS), 2020.
[8] Y. Zhang et al., “Theory of Mind for Multi-Agent Collaboration via Large Language Models,” arXiv preprint, 2023.
[9] C. Huyen, “AI Engineering: Building Applications with Foundation Models,” O’Reilly Media, 2024.
[10] S. Few, “Information Dashboard Design: Displaying Data for At-a-Glance Monitoring,” Analytics Press, 2013.
[11] [11] Knaflic, S. N., “Storytelling with Data: A Data Visualization Guide for Business Professionals,” Wiley, 2015.
[12] G. Box, G. Jenkins, G. Reinsel, “Time Series Analysis: Forecasting and Control,” Wiley, 2015.
[13] Hyndman, R. J., Athanasopoulos, G., “Forecasting: Principles and Practice,” OTexts, 2018.
[14] E. Turban et al., “Decision Support and Business Intelligence Systems,” Pearson, 2011.
[15] Gao, Y. et al., “Retrieval-Augmented Generation for Large Language Models: A Survey,” arXiv, 2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY