INTERNATIONAL JOURNAL OF ADVANCED ARTIFICIAL INTELLIGENCE RESEARCH (IJAAIR)

eISSN: 3087-4270

Volume. 02, Issue. 07, pp. 09-24, July 2025

EXPLAINABLE MACHINE LEARNING FOR FINANCIAL ANALYSIS

Olabayoji Oluwatofunmi Oladepo.
Department of Computer Science, Swansea University, Singleton Park, Sketty, Swansea, SA2 8PP, Wales, United Kingdom

Opeyemi Eebru Alao
Department of Computer Science, Swansea University, Singleton Park, Sketty, Swansea, SA2 8PP, Wales, United Kingdom

Article received: 22/06/2024, Article Accepted: 25/07/2025, Article Published: 29/07/2025
DOI: https://doi.org/10.55640/ijaair-v02i07-02

ABSTRACT

Technical analysis plays a vital role in the dynamic world of stock trading by helping traders identify patterns and trends in market prices. However, interpreting these patterns can be complicated, especially for novice traders, due to the complexity of charts and indicators involved. This research introduces a user-friendly tool aimed at making technical analysis more accessible. It provides a streamlined interface for analysing and forecasting stock trends, catering to users of varying expertise. Most existing trading platforms either lack predictive features or operate as opaque 'black-box' systems, which can deter beginner users. To address this, the research developed a solution that simplifies financial analysis and enhances clarity. This tool stands out by reducing reliance on complex visuals, improving user understanding of financial indicators, and explaining how they influence market predictions. A unique feature of the tool is its use of Explainable Artificial Intelligence (XAI), which provides transparency and builds user trust. Feedback from user testing was positive, with participants highlighting the tool's clarity and interactivity as valuable for making informed investment decisions.

KEYWORDS

Explainable AI, stock, support vector machine (SVM), random forest, rule builder.

INTRODUCTION

The rapid integration of machine learning (ML) into financial systems has revolutionized areas like credit scoring, fraud detection, algorithmic trading, and risk management (Jain et al., 2025). By leveraging complex algorithms from ensemble methods to deep neural networks financial institutions achieve unprecedented predictive accuracy and operational efficiency. However, this progress comes with a critical challenge: the opacity of "black-box" models. As ML systems increasingly influence high-stakes financial decisions, their lack of transparency poses significant risks, including regulatory noncompliance, eroded stakeholder trust, and unexplainable failures with systemic consequences (Mayenberger, 2019; Rane et al., 2023).

In the finance industry, trading involves the buying and selling of assets such as stocks, commodities, currencies, and bonds (Hayes, 2023). A critical component of successful trading is financial analysis, which involves evaluating financial data and market trends to guide investment decisions. This analysis generally falls into two categories: fundamental analysis, which examines economic and financial factors to assess a company's value, and technical analysis, which assumes that market patterns repeat and uses price and volume data to forecast future movements (Anbalagan and Maheswari, 2015; Epstein et al., 2017). The rise of digital trading platforms and low-cost brokerage services has made trading more accessible, attracting a wider audience. However, the inherent volatility of markets presents significant risks, particularly for individuals without a strong financial background. This highlights the need for intuitive and beginner-friendly tools (Epstein et al., 2017). Machine Learning (ML) models such as Support Vector Machines (SVM) and Random Forests have brought advanced forecasting capabilities to financial markets (Yang et al., 2002; Kane et al., 2014). Yet, many tools based on these models are hard to use due to their reliance on programming knowledge and often lack transparency, which can make users hesitant to trust them. To overcome these issues, this research work introduces a new analysis tool that blends traditional technical analysis with modern features. It leverages a Random Forest model to predict stock trends and integrates XAI techniques to make its decision-making process understandable. The tool also includes a Rule Builder that allows users to experiment with financial indicators and see how they influence predictions. This enhances learning and empowers users to engage more confidently with financial data.

1.0 METHODOLOGY

This section outlines the structured approach taken to develop the financial analysis tool. Because the research combines technical development, machine learning modelling, and user-centred evaluation, a mixed-methods strategy was adopted. This ensures both the robustness of the system and the relevance of the tool to end users.

1.1 Research Design

The research process followed a clear architecture, depicted in a project pipeline diagram (Figure 1). This pipeline outlines each stage, from data acquisition to model training, and finally, user evaluation.

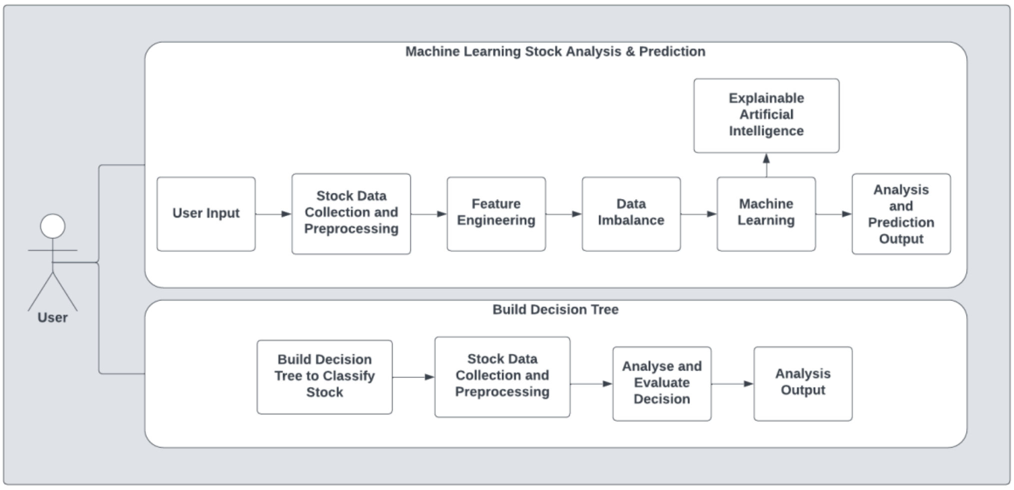


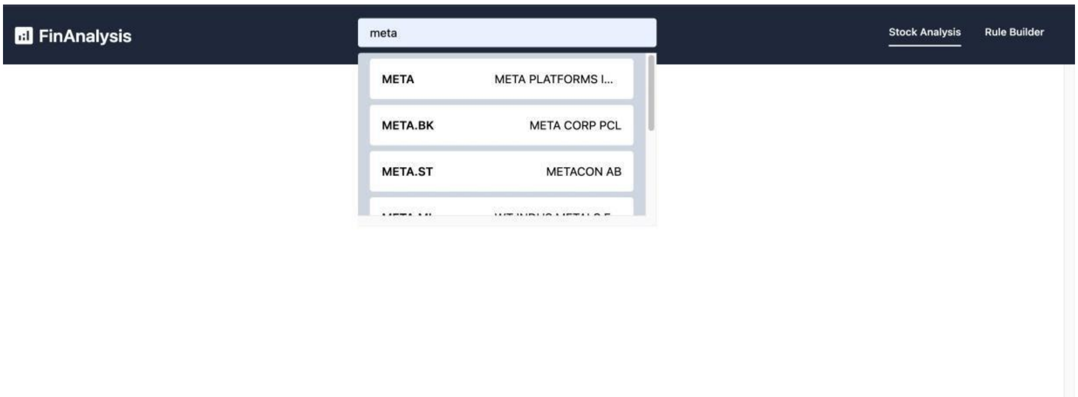
Figure 1: Research Pipeline

1.2 Machine Learning for Stock Trend Prediction

This section discusses the journey to harness machine learning for stock analysis, detailing every vital step in achieving this feat.

1.2.1 Data Collection & Preprocessing

Collecting and preparing data is crucial for building a reliable machine learning model, especially in finance, where markets are volatile. Data is sourced through the frontend of the tool, where users search for a stock ticker. Once selected, the backend fetches historical stock data using the `yfinance` Python library, which accesses Yahoo Finance's vast dataset. This data is also visualised with technical charts to aid user analysis.



(a) User Searching For A Specific Stock Ticker

(b) Financial Data Visualised Using Technical Chart

Figure 2: Visual Representations Of The Frontend Interface

After data retrieval, basic cleaning is performed, mainly dealing with missing values, assuming the base data is of high quality. Then, the user defines key parameters like trend threshold and prediction window (timeout), which are used to shape the predictive analysis. For instance, a timeout of 3 means the model will predict a trend three days into the future. To smooth out data and reduce noise, Exponential Moving Average (EMA) is applied. After this, the data is labelled based on trends—upward (1), downward (-1), or no trend (0)—using the user-defined parameters. These labels form the target variable for the classification task.

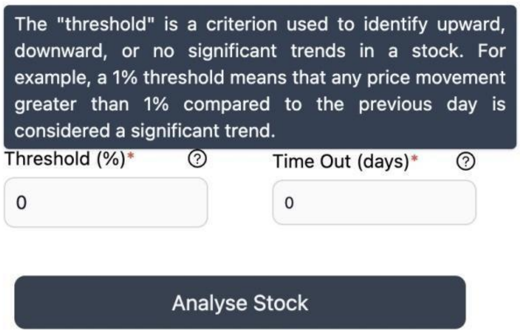


Figure 3: User Input Required For Pre-Processing

1.2.2 Feature Engineering

To build a strong model, meaningful features are engineered from financial indicators:

Simple Moving Average (SMA): The Random Forest model employs two intervals: a short 50-day window (SMA_50) and a long 200-day window (SMA_200). Additionally, an SMA_Crossover feature is formulated to identify the intersections between SMA_50 and SMA_200. This feature captures the crossover strategy, which implies that the asset acquires momentum when SMA_50 is above SMA_200 and loses it when below (Achelis, 2014). The SMA_Crossover is assigned a value of 1 when SMA_50 > SMA_200 and 0 when SMA_50 < SMA_200.

Moving Average Convergence Divergence (MACD): In engineering, the MACD feature, the MACD line and the Signal line are calculated, and a MACD_Crossover feature is introduced to capture the occurrences of crossovers. Notably, if the MACD line surges above the Signal line, MACD_Crossover is set to 1, symbolising a bullish signal; if it plunges below, it is set to -1, indicating a bearish signal (Mitchell, 2022; Hayes, 2021). The neutral state is represented by 0.

Bollinger Bands: A 20-day window and a standard deviation of 2 are used in the computation of the Bollinger Band indicator.

Relative Strength Index (RSI): An RSI above 70 signifies overbought conditions, while an RSI below 30 indicates oversold conditions (Epstein et al., 2017). The RSI status feature encapsulates this information, with values set to 1 for overbought, -1 for oversold, and 0 for neutral states.

Stochastic Oscillator: The stochastic status feature is used to identify overbought and oversold conditions. A stochastic status value of 1 signifies overbought conditions, -1 represents oversold, and 0 designates a neutral state.

By harnessing these diverse technical indicators, the model gains the capability to analyse multiple aspects of market dynamics, thereby amplifying its predictive accuracy.

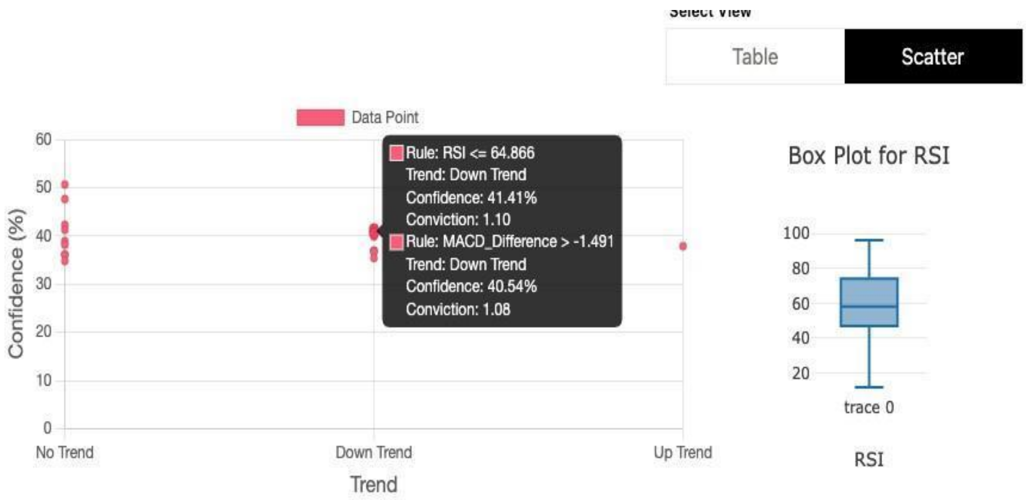


Figure 4: A combined visual representation illustrating the model’s intricate decision dynamics: (Left) A scatter plot provides a view of the correlation between the model’s confidence in each rule and the anticipated trend; (Right) A box plot showing the range and distribution of values considered by the Random Forest. Other features include: OBV (On-Balance Volume), Mass Index, Williams %R and Rate of Change (ROC).

These features help the model understand multiple market dimensions such as trend, momentum, and volatility, thus improving prediction accuracy.

1.3 Machine Learning Model

A major objective of this research is to develop a reliable machine learning model capable of forecasting stock market trends. The selected algorithm is Random Forest, known for its ability to handle large, imbalanced datasets and perform well in high-dimensional spaces. Additionally, Random Forest has a built-in mechanism for identifying which features are most important, making it ideal for integration with Explainable AI (XAI). The model was trained as a classification problem using a Random Forest Classifier, with its performance optimised by grid search to fine-tune the hyperparameters. This ensures a good balance between bias and variance while maintaining efficiency. One of the common challenges in stock prediction is class imbalance, which is that periods of market stability far outnumber upward or downward trends. To address this, the SMOTE (Synthetic Minority Over-sampling Technique)algorithm was used. SMOTE creates synthetic samples for underrepresented trend classes, improving the model's ability to learn from all types of data. The dataset was then split into features (e.g., technical indicators) and a target variable (the trend classification). An 80/20 train-test split was applied, which is a standard practice in classification tasks.

To evaluate the model, several metrics were used:

F1 Score (Weighted): Balances precision and recall across all classes.

Accuracy: Measures the overall correctness of the model’s predictions.

Macro Precision: Averages precision scores across all classes, ensuring fair treatment of imbalanced data.

Once trained and evaluated, the model's predictions and performance metrics are presented to the user, giving them immediate feedback on the tool's reliability.

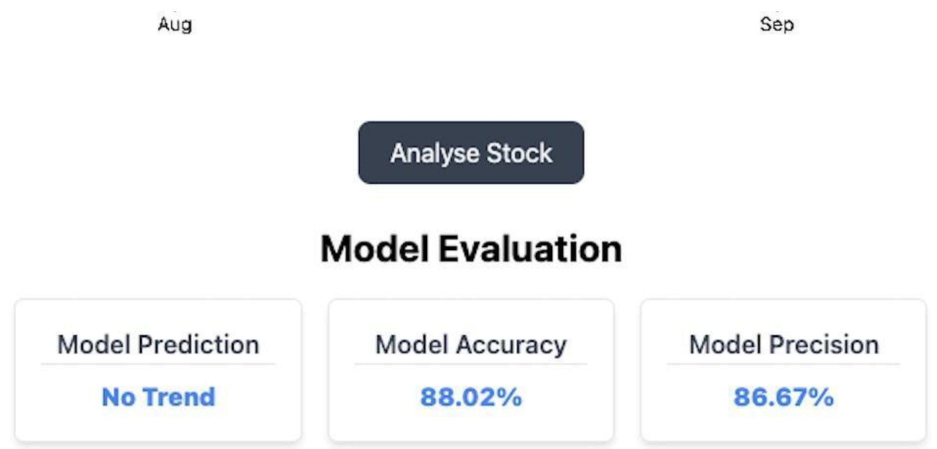


Figure 5: Model Prediction And Evaluation

1.4 Explainable AI (XAI) Implementation

A key part of the research is making sure users understand how and why the model makes predictions. To do this, the tool integrates explainability using both built-in feature importance and a novel rule-based method inspired by the work of Hatwell, Gaber, and Azad.

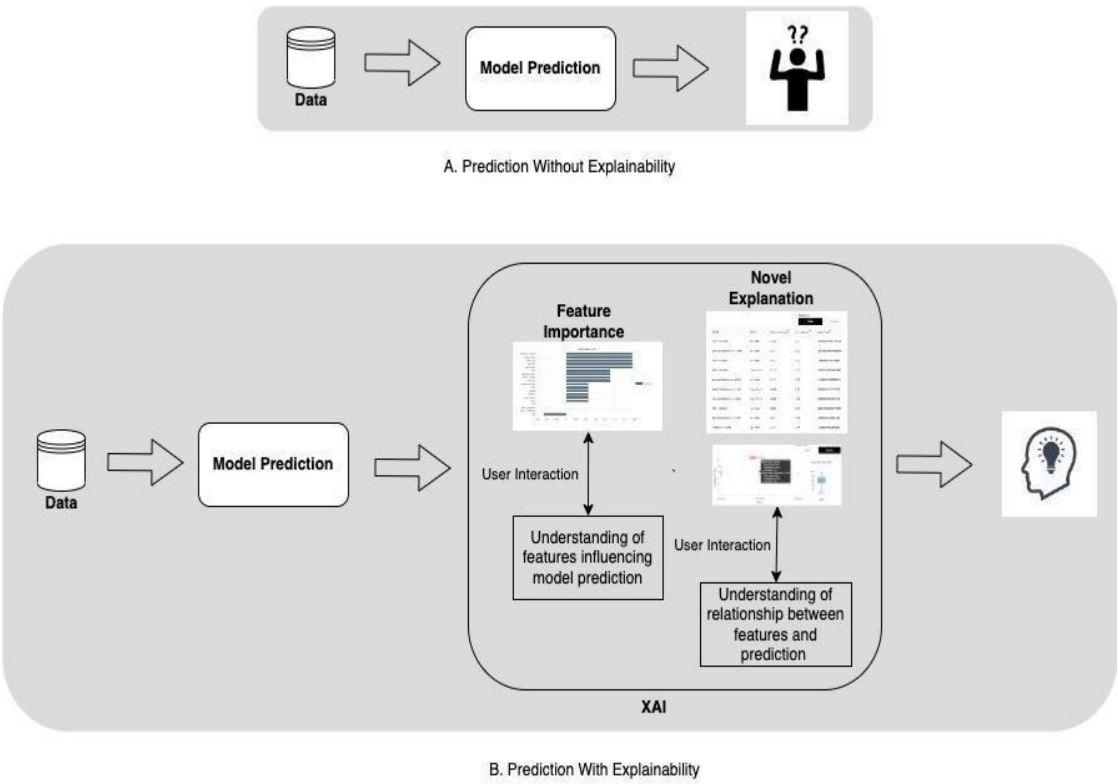


Figure 6: Illustration Of The Benefit Of Explainable AI Prediction

1.4.1 Feature Importance Visualisation

One advantage of using Random Forest is that it automatically calculates the importance of each feature (i.e., technical indicator). This is visualised through a **bar chart**, allowing users to see which indicators most influenced the prediction. This visual explanation increases transparency and helps users build trust in the tool's output. For enhanced user accessibility and interaction, this feature importance metric is visualised using a bar chart, as shown in Figure 17.

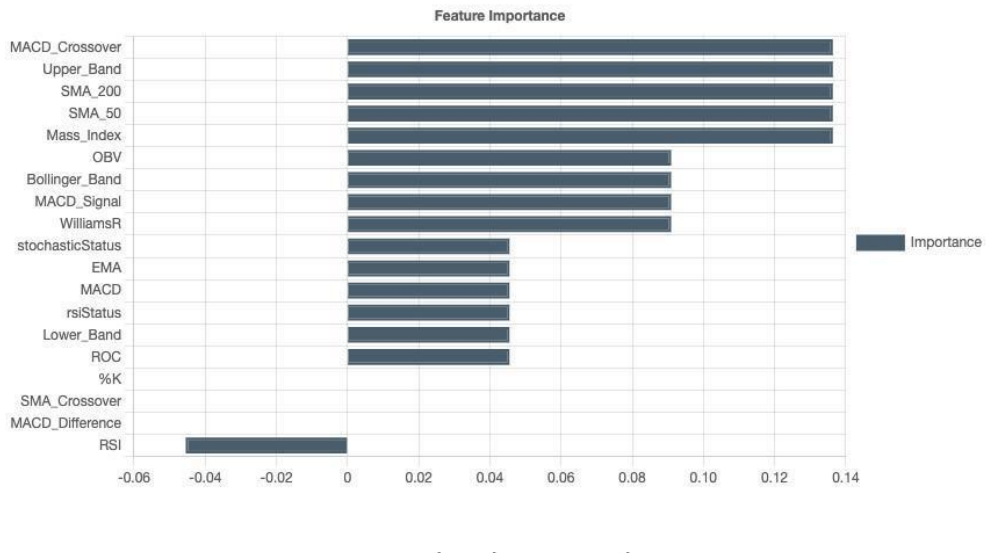


Figure 7: Features (Technical Indicators) Ranked According To Their Importance

1.4.2 Rule Extraction from Random Forest

Beyond feature ranking, a deeper form of explainability is provided by extracting decision rules from the Random Forest. Each decision tree in the model holds a set of rules that describe how predictions are made. These rules are:

1. Extracted by traversing each tree from root to leaf.

2. Formatted into a simplified, human-readable version using custom Python functions.

3. Analysed using the Apriori algorithm to find frequent patterns—combinations of feature conditions that commonly appear in successful predictions.

This process reveals how certain feature combinations influence the model's predictions, offering users a clear look into the model's internal logic.

1.4.3 Decision Tree Builder (Rule Builder)

The Decision Tree Builder, also known as the Rule Builder, is a standout feature of the tool. It offers users an interactive way to design, test, and evaluate custom strategies for classifying stock trends. The main goals of this feature are to:

1. Let users construct their own decision trees for stock prediction.

2. Help users incorporate insights gained from the Random Forest model into their custom trees.

This tool is especially beneficial for users who want to combine human intuition with machine-derived logic. The interface is intuitive and designed for both beginners and experts. Users can select specific features (e.g., RSI, MACD crossovers) and define what happens when those conditions are met or unmet, creating a full decision path. Behind the scenes, this feature uses code (provided in the Appendix) to build and process these user-defined decision trees. Users can also refer to the patterns extracted from the Random Forest to guide their decision-making, effectively blending explainable AI with personalised strategy building. In essence, the Rule Builder acts as both a learning tool and a strategy tester, helping users better understand the relationships between financial indicators and their influence on stock trends.

3.4.4 Traditional Charting Feature

Recognising the importance of traditional charting in technical analysis, the tool includes a classic charting module to complement its machine learning capabilities. Many traders are already familiar with traditional charts, so this feature helps bridge the gap between conventional analysis and modern AI-driven insights. These charts display well-known technical indicators such as Moving Averages, RSI, MACD, Bollinger Bands, and more. The design ensures that both beginner and experienced users can easily interpret the visuals and compare them with the tool's predictions. By offering both machine-generated predictions and traditional visual tools, the system caters to a wider range of user preferences and trading styles.

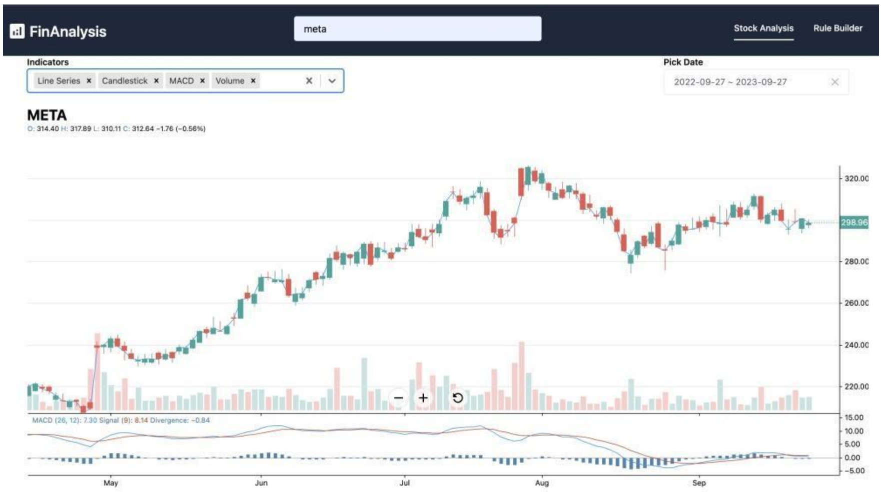


Figure 8: Traditional Charting Feature

1.5 Software Development

To manage the development of the tool efficiently, the team adopted the Agile methodology, specifically using the Kanban approach. Tasks were organised using Trello, a visual project management tool. Each task was represented by a card and moved through stages such as "To Do," "In Progress," and "Completed." This method helped track progress and ensure the timely delivery of research milestones. A two-week buffer was also added to the timeline to accommodate any unforeseen delays (see Figure 9).

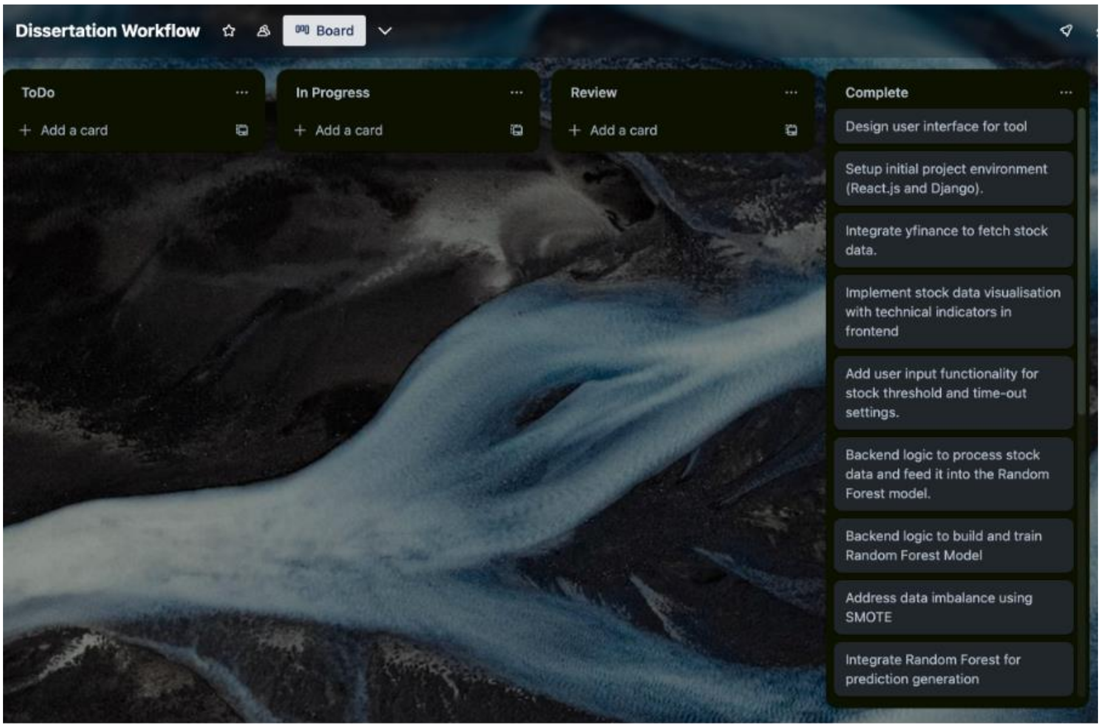


Figure 9: Kanban Board Showing Task Status After Research Completion

In addition to Kanban, a Gantt chart was used to visualise the overall timeline, allocate resources effectively, and monitor progress against deadlines (see Figure 10).

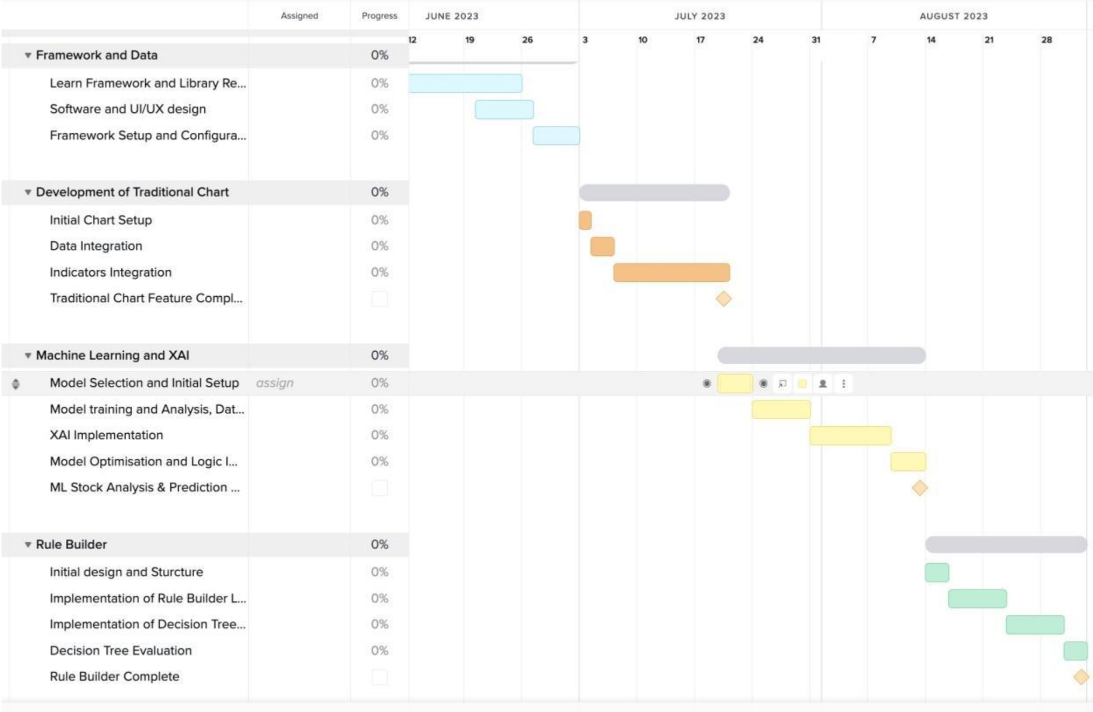


Figure 10: Gantt Chart of Research Schedule

2.0 Implementation and Evaluation

2.1 Development Challenges

The development process encountered several challenges:

2.1.1 Deployment Issues

Although the tool worked perfectly in a local environment, deploying it online presented problems primarily due to the long training and testing times of the machine learning model, which led to timeouts. This was resolved by integrating DjangoQ, an asynchronous task queue, allowing background processing.

2.1.2 Charting Difficulties

Implementing stock data visualisation with the `react-financial-chart` library was problematic due to limited documentation. These hurdles were overcome through community support and iterative trial-and-error.

2.2 Risk Management

Risks were assessed using a structured formula where Importance (I) = Chance (C) × Severity (S). The team identified and addressed multiple potential risks, including:

Delays in research completion

Challenges with chart and indicator implementation

Difficulties in model deployment or algorithm performance

Complexity in building the Rule Builder

Lack of user understanding of complex ML models

Unfamiliarity with third-party libraries

Risk of data loss

Each risk was paired with mitigation strategies like early prototyping, regular backups, using reliable libraries, and extensive research. For example, data loss was managed using cloud storage (e.g., OneDrive), GitHub version control, and regular integrity checks. See Table 1 in the original document for a full breakdown.

Risk	C	S	I	Mitigation
Delays in Research Completion	3	7	210	Adopt the Kanban method to streamline workflow,

				ensuring timely completion. A buffer of 2 weeks added to the research timeline to account for unforeseen delays.
Implementation of traditional charts and indicators	6	70	420	Comprehensive research on available libraries and techniques. Familiarisation with stock market basics to ensure accurate representations.
Failure in implementing ma- chine learning algorithms	4	85	340	Leveraging existing libraries like scikit-learn. Setting aside dedicated time for research and learning.
Difficulty in explaining complex ML models to users	6	70	420	Adopting comprehensive XAI techniques, virtual workshop, and intuitive visualisations.
Challenges in building a response, user-friendly interface	6	50	300	Using a reliable front-end frame- works, continuous user feedback loops, and iterative UI improvements.

Shortcomings in utilising exter- nal libraries	3	75	225	Deep diving into the library documentation, exploring community support forums, and having backup libraries in mind.
Complexity implementing deci- sion tree builder and logic	6	85	510	Preliminary prototyping to understand the challenges better and allocation of extra develop- ment time to cater to logic complexity.
Loss of research data or corruption	5	100	500	Routine backups using the cloud platforms (OneDrive), GitHub for version control, and data integrity checks.

Table 1: Risk Assessment and Mitigation Strategies

2.3 Software Testing

To ensure the tool's reliability, a thorough testing process was followed. During development, ad hoc testing was used to catch unexpected issues. This informal approach uncovered a range of bugs from minor UI problems to major logic flaws in the ML module. One specific bug involved the model returning incomplete predictions. This was traced, corrected, and followed by a review of related components to prevent similar issues. The testing process played a crucial role in refining the tool and ensuring it functioned correctly under different conditions.

2.4 Evaluation

This section evaluates the effectiveness of the developed tool, focusing on how users perceived its functionality, ease of use, and the impact of its explainability features.

2.4.1 Workshop Setup

To prepare participants for the user study, a virtual workshop was organised. It introduced 14 participants to essential concepts in financial analysis and machine learning. Their backgrounds ranged from no prior trading experience to moderate knowledge Figure 11).

Figure 11: Distribution of Trading Knowledge Among Study Participants

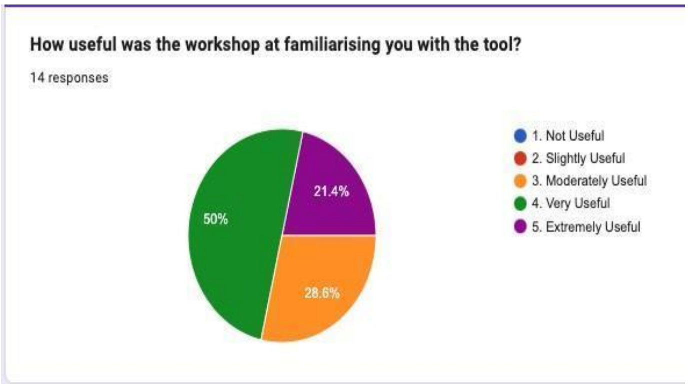


Figure 12: Effectiveness of Workshop

The workshop served as a crash course on:

- Basic trading concepts
- How the tool works
- Key features like prediction, explainability, and rule-building

Participants rated the workshop highly, with an average score of 4.0 out of 5, indicating it effectively prepared them to use the tool (Figure 12).

2.4.2 User Study

After the workshop, participants were asked to use the tool and complete a survey. The survey aimed to measure:

- Tool effectiveness
- User trust in machine learning predictions
- The value of Explainable AI (XAI)

2.4.3 Ease of Use

Usability was a key concern, especially since the tool targets users with varying levels of expertise. Results were positive:

- The average rating for ease of use was 4.14 out of 5 (see Figure 13).
- Participants praised the clean interface and intuitive design.
- One participant said it was "very user-friendly and suitable for amateur traders."

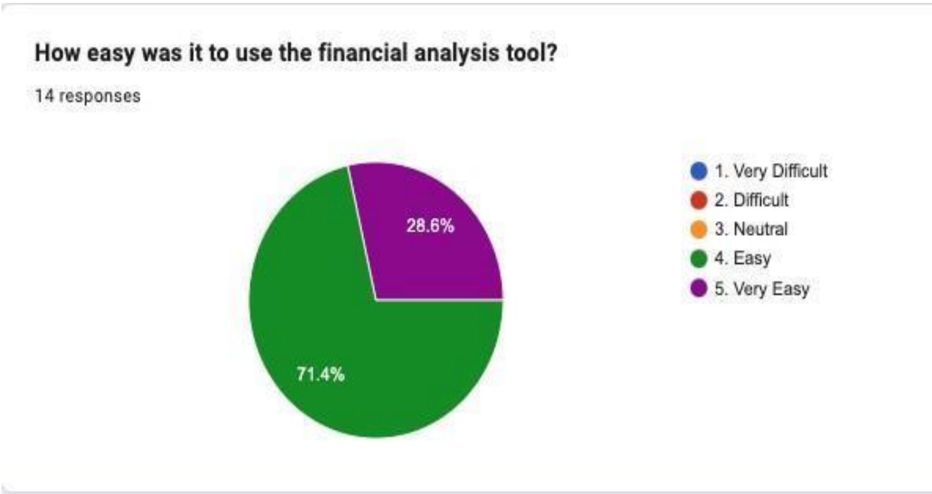
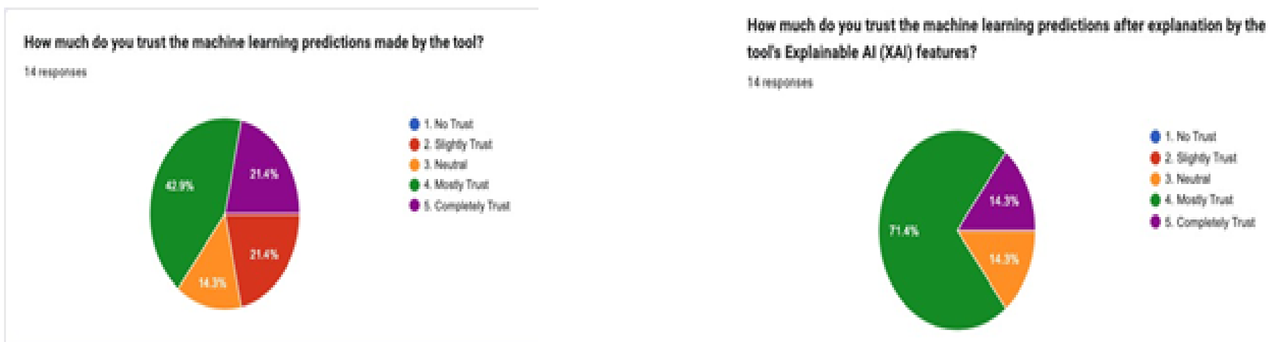


Figure 13: User Study Response To Ease Of Use Of The Proposed Tool

This feedback suggests the tool lowers the barrier for beginners to participate in stock trading and analysis confidently.

2.5 Trust in Machine Learning and the Role of XAI

Trust plays a major role in whether users adopt AI-driven tools. Before being introduced to the tool's explainability features, participants reported moderate trust in the predictions (average score: 3.64/5). After using the XAI features, trust levels rose to 4.0/5 (Figure 14). This increase shows that understanding how predictions are made helped users feel more confident in the tool.



(a) Trust Distribution Without XAI

(b) Trust Distribution With XAI

Figure 14: Comparison of User Trust in Machine Learning Predictions: Before and After Incorporation of XAI

Additional insights:

- Participants appreciated the **extracted rules from the Random Forest**, particularly when used with the Rule Builder.
- One user noted these rules were useful enough to stand alone as a feature.

Participants were also asked if the extracted rules helped them build better strategies. The average usefulness rating was 3.5/5 (Figure 15), showing that explainability not only improved trust but also supported more informed decision-making.

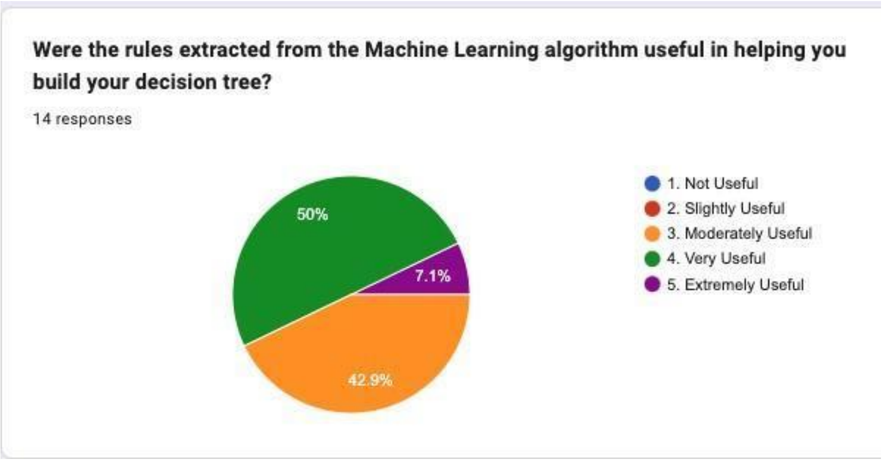


Figure 15: User Feedback on the Usefulness of Extracted Rules

2.6 Rule Builder Feedback

The Rule Builder was widely seen as the most valuable feature (Figure 15):

- Over half of the participants identified it as their favourite.
- The average usefulness rating was 3.93/5 (Figure 16).
- Users appreciated the interactive, hands-on approach to learning and strategy building.
- One called it a "brilliant feature" for allowing personalised strategy creation.

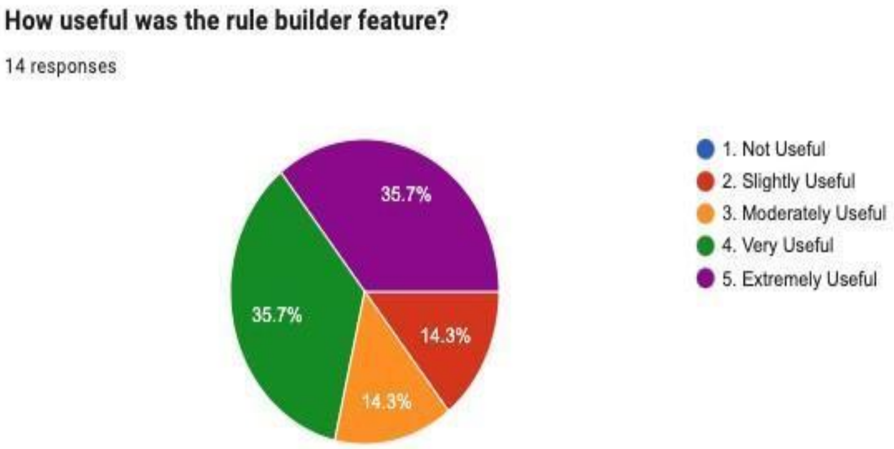


Figure 16: User Feedback on the Usefulness of the Rule Builder Feature

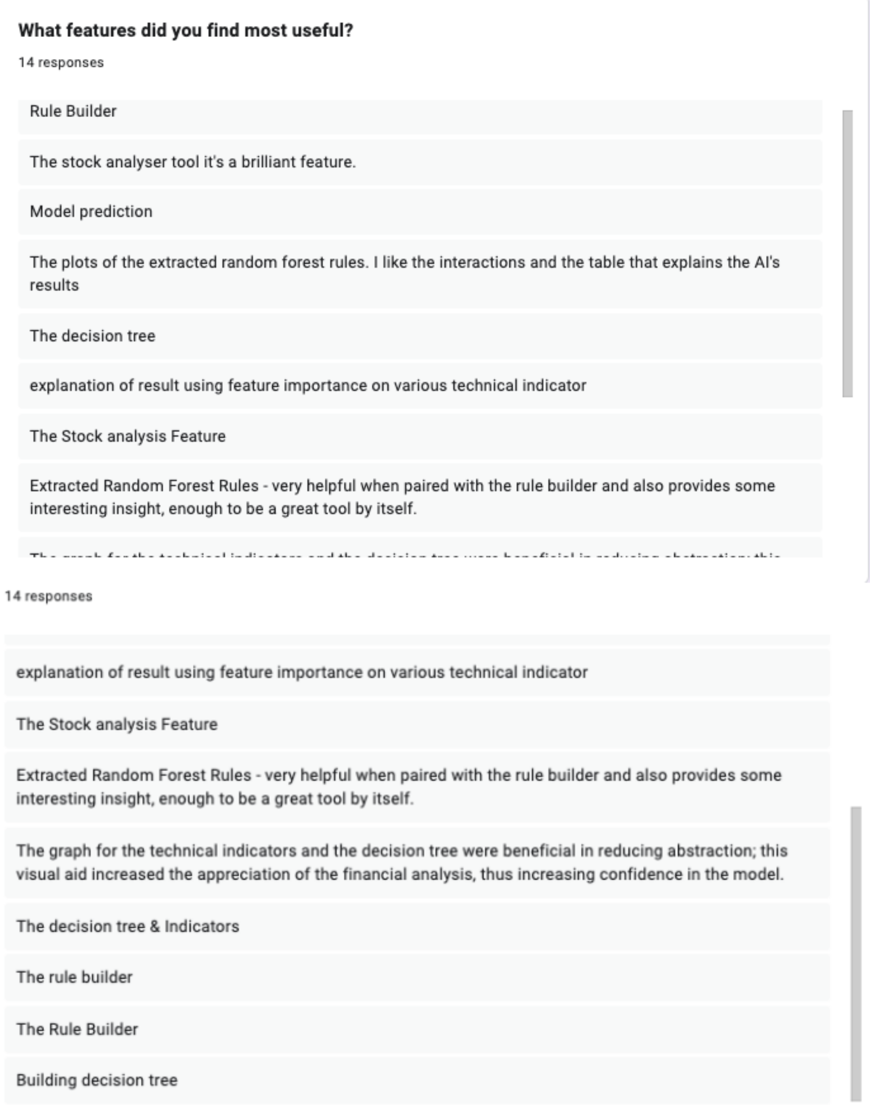


Figure 17: User Responses To Most Useful Feature

The feedback confirmed that the Rule Builder helped users visualise how different financial indicators relate to stock trends, making complex financial concepts more accessible.

3.0 CONCLUSION

This research, although challenging, has been immensely rewarding, offering valuable insights into the complexities of financial trading and the potential of technology to make it more accessible and understandable for everyone. This research has successfully developed an innovative financial analysis tool designed to deconstruct the intricate landscape of stock trading. Reflecting on this research's objectives, it is satisfying to acknowledge that each objective has been successfully met. The main objectives of this research include the development of a cutting-edge tool leveraging a Random Forest model to predict stock trends, the integration of XAI into the tool to provide transparency and foster trust, and the provision of an intuitive interface enabling users to construct a decision tree to predict stock trends. A pivotal contribution of this research has been its focus on bridging the existing trust gap between users and AI-based predictive tools. The implementation of XAI has elevated the user experience by offering transparency and insight into the decision-making processes of the AI, enabling users to understand and trust the predictions made by the tool better. In our user study, trust levels in the tool's predictions increased by 9.89% on average after traders interacted with the explanation interface, validating that the tool has empowered users, particularly those new to stock trading, to make more informed and confident trading decisions. Another noteworthy contribution achieved during this research is the incorporation of a user-centric feature allowing users to construct their decision trees, which has proven to be instrumental in enhancing user understanding of the relationships between various financial indicators and their implications on stock trends. In conclusion, this research has unveiled the transformative potential of integrating user-centric design, advanced machine learning models, and transparent AI in financial trading tools. It reveals how the intricate realm of stock trading can be made approachable and comprehensible to all, regardless of financial knowledge.

3.1 Future perspective

While the research has achieved its objectives, the scope for future advancements and refinements remains expansive. Below are several potential avenues for future development:

1. Enhanced Learning Resources: Introducing more comprehensive learning resources and tutorials within the tool can help users, especially beginners, better understand financial indicators and machine learning concepts, fostering an environment of learning and exploration.

2. Selective Decision Tree Loading and Modification: A significant enhancement would be to allow users to selectively load a decision tree from the Random Forest, modify the loaded tree using the Decision Tree Builder and test how well the tree performs at classifying trends as illustrated in Figure 18. This feature can empower users to delve deeper into the model's rationale and tweak it according to their preferences and insights, promoting a more personalised strategy and user experience.

3. Enhanced Decision Path Extraction: The current algorithm used to extract decision paths from the random forest model extracts all paths irrespective of whether they lead to correct or incorrect predictions. An area of future work is enhancing this algorithm to only extract decision paths that result in accurate predictions on the training data. By focusing exclusively on the "right" decision paths, the extracted rules are more likely to produce reliable predictions when new data is fed through them. This creates a filtered set of high-quality rules that users can inspect and experiment with, without the noise from low-performing rules.

4. Comprehensive User Study: While the initial user study provided valuable insights, more extensive evaluations are needed to fully gauge the tool's benefits for traders. Future work could conduct large-scale user studies with more participants across diverse demographics and trading experience levels. This would shed light on long-term usage patterns and how trust and understanding evolve with increasing familiarity with the tool. Additionally, the impact of the extracted Random Forest rules on users' personalised strategies can be evaluated. Participants can be asked to build their own decision tree prior to viewing the extracted rules. The performance of their decision tree in predicting stock trends can then be compared to a second simulation where they incorporate insights from the rules.

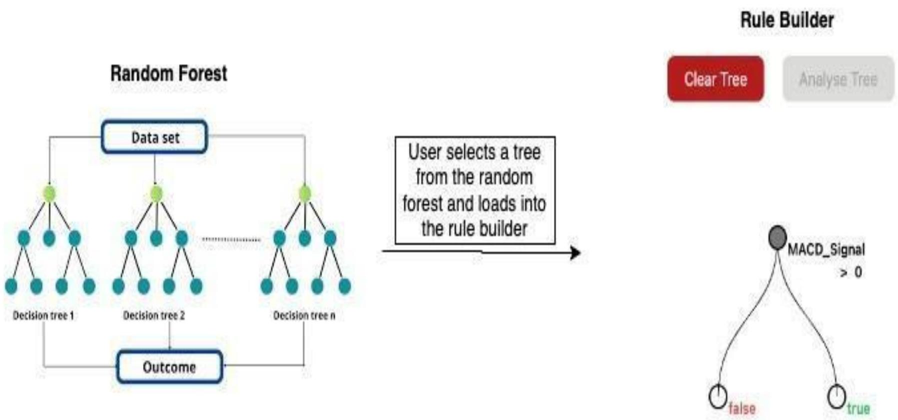


Figure 18: Illustration Of The Proposed Selective Decision Tree Loading And Modification Feature

REFERENCES

1. Achelis, S. B., & Achelis, S. (2001). Technical analysis from A to Z (pp. 199–200). New York: McGraw-Hill.

2. Anbalagan, T., & Maheswari, S. U. (2015). Classification and prediction of stock market index based on fuzzy metagraph. Procedia Computer Science, 47, 214–221.

3. Epstein, L., & Roze, G. D. (2023). Trading for dummies. John Wiley & Sons.

4. Hayes, A. (2021). Stochastic oscillator: What it is, how it works, how to calculate. Investopedia. URL: https://www.investopedia.com/terms/s/stochasticoscillator.asp (Accessed: October 16, 2024).

5. Hayes, A. (2022). Financial markets: Role in the economy, importance, types, and examples. Investopedia, 31(5), 34–43.

6. Jain, D. K., Bhatele, K. R., Garg, D., & Dhiman, G. (Eds.). (2025). A Compendium of Responsible Artificial Intelligence. CRC Press.

7. Kane, M. J., Price, N., Scotch, M., & Rabinowitz, P.

(2014). Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. BMC Bioinformatics, 15, 1–9.

8. Mayenberger, D. (2019). How to overcome modelling and model risk management challenges with artificial intelligence and machine learning. Journal of risk management in financial institutions, 12(3), 241-255.

9. Rane, N., Choudhary, S., & Rane, J. (2023). Explainable Artificial Intelligence (XAI) approaches for transparency and accountability in financial decision-making. Available at SSRN 4640316.

10. Yang, H., Chan, L., & King, I. (2002, August). Support vector machine regression for volatile stock market prediction. In International conference on intelligent data engineering and automated learning (pp. 391–396). Berlin, Heidelberg: Springer Berlin Heidelberg.