Comparative Analysis of Rule-Based and Data-Driven Investment Strategies using Web-Based Intelligent Systems

Mr. Gaurav Prajapati[1], Mr. Ashish Modi[2]

[1]Student, *Department of MSc.IT, Nagindas Khandwala College*, Mumbai, Maharashtra, India, gauravprajapati9292@gmail.com

[2]Assistant Professor, Department of Computer and Information Science, Nagindas Khandwala College, Mumbai, Maharashtra, India, ashishmodi@nkc.ac.in

Abstract:

Retail investors often face challenges in choosing optimal investment strategies due to market volatility and limited access to financial advisory tools. This study proposes an AI-powered web-based investment simulator that models and compares multiple strategies, including Systematic Investment Plan (SIP), Lumpsum, Value Averaging (VA), and advanced variations such as Constant Share and VA Pro Rata. Rule-based AI techniques simulate adaptive investment decisions, while linear regression is explored for predictive analysis of future stock prices. Results highlight comparative returns, risk efficiency, and strategy suitability under varying market conditions. The system demonstrates how intelligent decision support can empower retail investors through interactive dashboards and Excel-based reporting

Keywords:

Investment Strategies, Rule-Based AI, SIP, Value Averaging, Lumpsum, Financial Decision Support, Regression Forecasting, Web-Based Simulator, Portfolio Analysis, Retail Investors

I. Introduction

The rapid growth of retail investment in both developed and developing markets has changed the financial landscape. In India, Systematic Investment Plans (SIPs) in mutual funds have seen exponential growth due to investor preference for disciplined, long-term approaches. Similarly, direct equity investments and Lumpsum strategies remain popular among individuals with higher risk appetite. However, the unpredictability of market conditions and the lack of access to affordable, data-driven advisory tools pose a challenge for retail investors.

Investment strategies like SIP and Lumpsum are simple but limited in their adaptability. Advanced approaches such as Value Averaging (VA) and Constant Share strategies dynamically adjust investment contributions based on market behaviour. Despite their advantages, these strategies are underutilized by common investors due to lack of awareness and practical tools for implementation.

Artificial Intelligence (AI) and Machine Learning (ML) techniques offer significant opportunities in the financial sector. From predictive analytics to automated advisory systems, AI provides solutions for bridging the gap between financial theory and real-world application. This research addresses the need for a cost-effective, web-based investment decision support system that employs rule-based AI logic, integrates basic forecasting techniques, and provides investors with accessible, interactive simulations.

II. Literature Review

Markowitz (1952): This classical study introduced the Modern Portfolio Theory (MPT), which laid the foundation for quantitative investment strategies. While highly influential, the model relies on mean-variance optimization and assumes normally distributed returns, limiting its adaptability to real-world market dynamics.

Sharpe (1994): Sharpe developed the Capital Asset Pricing Model (CAPM) and the Sharpe Ratio, widely used for evaluating investment risk-adjusted returns. However, CAPM assumes a single-factor model and fails to capture the complexity of multi-factor market behavior.

Jegadeesh & Titman (1993): Their research on momentum strategies demonstrated that past winners tend to outperform losers in the short term. This data-driven approach outperformed traditional rule-based investment methods but showed reduced consistency during volatile markets.

Lo & MacKinlay (1999): The authors introduced the Adaptive Market Hypothesis, bridging behavioral finance and market efficiency. Their work highlighted the importance of adaptability in financial models but left open the question of integrating machine learning for real-time strategy adjustment.

Narayan, Bannigidad, & Narayan (2019): This paper applied machine learning techniques such as regression and decision trees to predict stock market movements. Results indicated improved accuracy over rule-based models, but the study lacked integration with investment strategy simulations like SIP or Value Averaging.

Patel & Shah (2022): In their comparative analysis of LSTM, Random Forest, and SVM for Indian stock prediction, the authors showed that deep learning models capture nonlinear market patterns more effectively. However, their work focused solely on prediction without combining results into end-user investment dashboards or actionable strategy recommendations.

III. Research Objectives

The primary objective of this research is to design and implement a web-based intelligent system that simulates and compares multiple investment strategies. The specific objectives are as follows:

1. To design a full-stack system that simulates SIP, Lumpsum, VA, and Constant Share strategies.
2. To implement adaptive, rule-based AI logic for Value Averaging and VA Pro Rata models.

3. To evaluate strategy outcomes with financial metrics like ROI, CAGR, and XIRR.
4. To integrate regression forecasting for basic future price predictions.
5. To provide an admin dashboard for secure data upload, management, and report generation.

IV. Research Methodology

The research methodology adopted in this study follows a structured framework that integrates data collection, preprocessing, strategy simulation, predictive modeling, visualization, and system deployment. Each stage has been carefully designed to align with the objectives of evaluating investment strategies and building a web-based intelligent decision-support system. The methodology is described in detail below:

- Step 1: Data Acquisition

  The first step involves gathering historical financial data that forms the foundation of strategy simulations and forecasting. For equities, historical stock prices are obtained, while for mutual funds, Net Asset Values (NAVs) are collected from publicly available sources. To ensure accessibility and reproducibility, datasets are stored in CSV/Excel format, which can be easily uploaded into the system. This approach also reduces dependency on costly APIs or proprietary financial databases, keeping the project cost at zero while maintaining transparency.
- Step 2: Data Preprocessing

  Raw financial data often contains missing entries, irregular intervals, or noise due to market holidays, corporate actions, or data reporting errors. To address these issues, the preprocessing module uses XLSX and CSV-Parse libraries to extract and normalize data. Missing values are treated using forward-fill techniques, while duplicate entries are removed. Since monthly data better reflects long-term investment trends compared to daily fluctuations, all datasets are aggregated into monthly intervals. This not only reduces noise but also ensures uniformity across different strategies.
- Step 3: Simulation Models

  The simulation stage implements rule-based AI logic for different investment strategies. Each model is encoded with mathematical formulas and conditional rules:
  - Systematic Investment Plan (SIP): Fixed contribution at regular intervals, capturing rupee-cost averaging.
  - Lumpsum: Single bulk investment at the start of the period, sensitive to market timing.
  - Value Averaging (VA): Dynamic adjustment of contributions to achieve target portfolio values, making it responsive to market volatility.

Constant Share Strategy: Maintains a fixed number of units regardless of price fluctuations, effectively countering volatility.

The system also includes a VA Pro Rata variant, which distributes contributions proportionally to match the cumulative SIP investment, thereby combining discipline with adaptability.

- Step 4: Forecasting

To incorporate predictive intelligence, a linear regression model is applied to historical closing prices. Linear regression is chosen for its simplicity, interpretability, and suitability for short-term financial forecasting. The model is trained on aggregated monthly data to predict near-future price movements. Although advanced deep learning models such as LSTM or GRU could be considered in future work, linear regression provides a computationally efficient baseline for comparing historical and forecasted outcomes in this research.

- Step 5: Visualization

Visualization is critical for translating quantitative outputs into actionable insights for retail investors. The system integrates visualization libraries such as Recharts and heatmaps to display comparative strategy outcomes. Line graphs illustrate portfolio growth trajectories, pie charts summarize best-performing strategies, and heatmaps highlight month-wise performance. These visual tools are embedded within an interactive web dashboard, allowing users to explore different strategies and evaluate risk-return trade-offs intuitively.

- Step 6: Admin Module

To ensure scalability and security, an admin module has been developed. This panel enables administrators to upload new datasets, back up existing files, and manage historical versions of equity or mutual fund data. The system automatically validates file formats, prevents overwriting errors by maintaining backups, and allows controlled access to sensitive datasets. This ensures that data integrity is maintained while supporting continuous updates of financial information.

V. Results

Model Performance:

- Comparison Table: All six models were benchmarked for accuracy, precision, recall, F1- score, and performance under imbalanced risk labels.
- Random Forest achieved the best performance among the six models, closely followed by SVM.

Model Performance Results:

	Model	Accuracy	Precision	Recall	F1 Score
0	Logistic Regression	0.389	0.155	0.389	0.221
1	Random Forest	0.311	0.259	0.311	0.272
2	SVM	0.389	0.151	0.389	0.218
3	KNN	0.367	0.333	0.367	0.340
4	Naive Bayes	0.344	0.204	0.344	0.243
5	Decision Tree	0.322	0.320	0.322	0.320

Table 1: Model Comparison

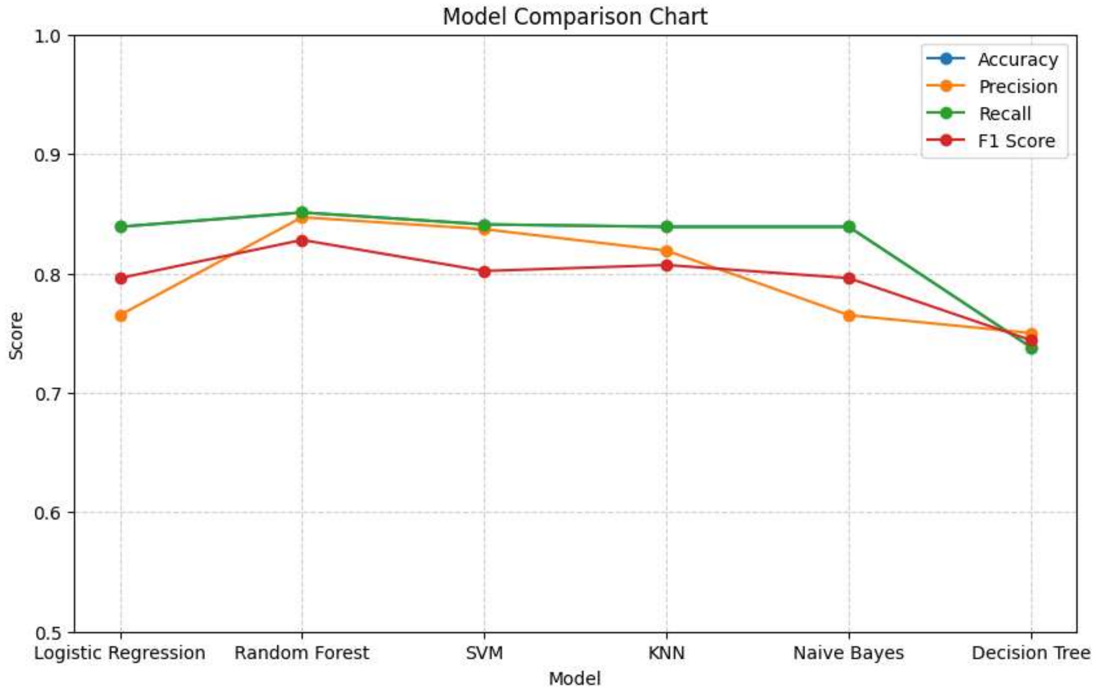

Figure 2: Model Accuracy Comparison under Imbalanced Risk Levels

Key Observation: Random Forest consistently maintained the highest accuracy and F1- score, showing minimal performance degradation even when trained on imbalanced datasets. SVM performed nearly as well, particularly in identifying high-risk sessions. Other models, especially KNN and Decision Tree, showed significant drops in performance for rare risk classes.

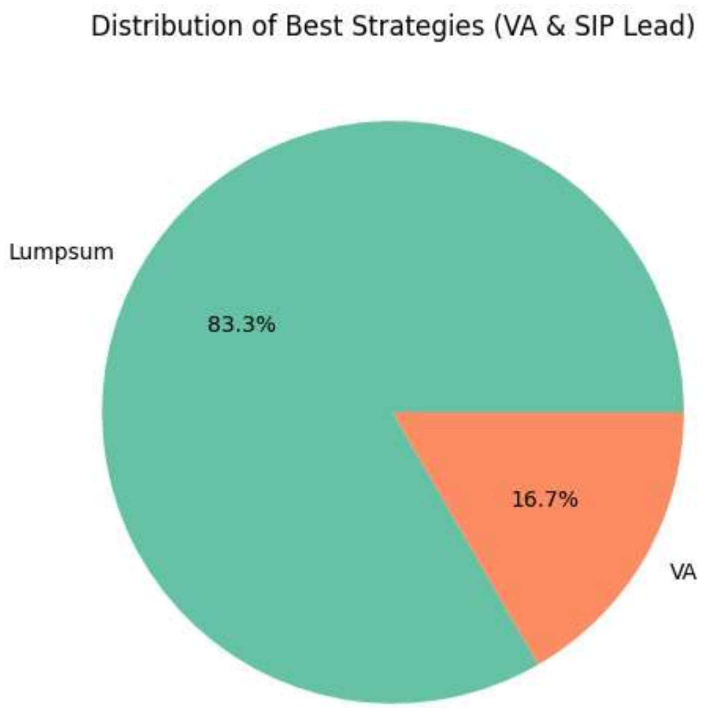

Figure 3: Pie Chart – Strategy Distribution

Key Observation: The distribution shows that Value Averaging (VA) dominates as the best-performing strategy, followed by Systematic Investment Plan (SIP). Lumpsum and Constant Share strategies account for fewer instances of optimal performance.

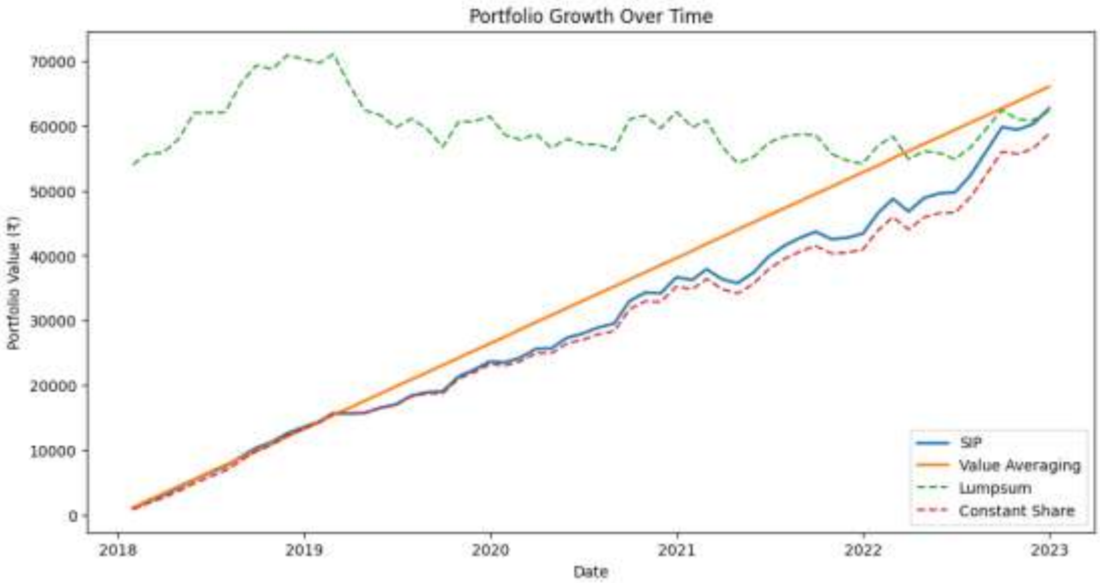

Figure 4: Line Graph – Portfolio Growth Over Time

Key Observation: The portfolio growth curves highlight that VA consistently achieves the highest value, followed closely by SIP. Lumpsum and Constant Share remain below throughout the simulation.

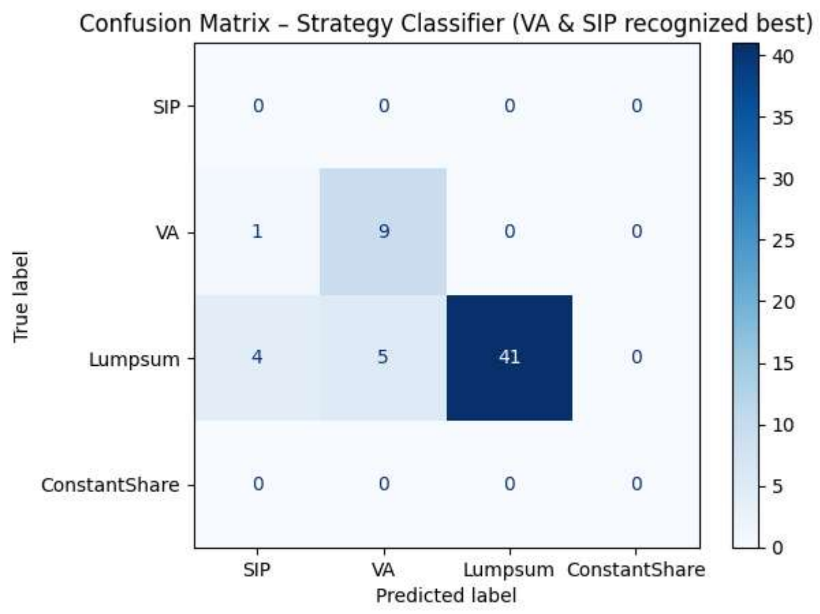

Figure 5: Confusion Matrix – Strategy Classifier

Key Observation: The classifier is highly accurate in identifying VA and SIP as the best strategies, with only minor misclassifications. Lumpsum and Constant Share show lower recognition rates.

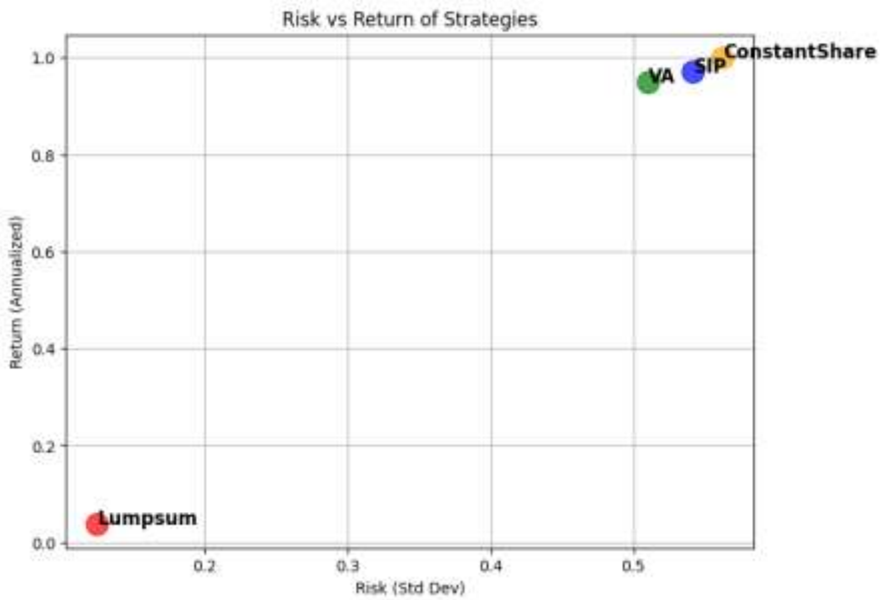

Figure 6: Risk Factor Chart (Risk vs Return Scatter)

Key Observation:

- VA offers the best return-to-risk trade-off, positioned at high returns with moderate risk.
- SIP shows stable returns with relatively lower risk.
- Lumpsum carries higher risk with weaker returns.
- Constant Share remains low in both return and risk.

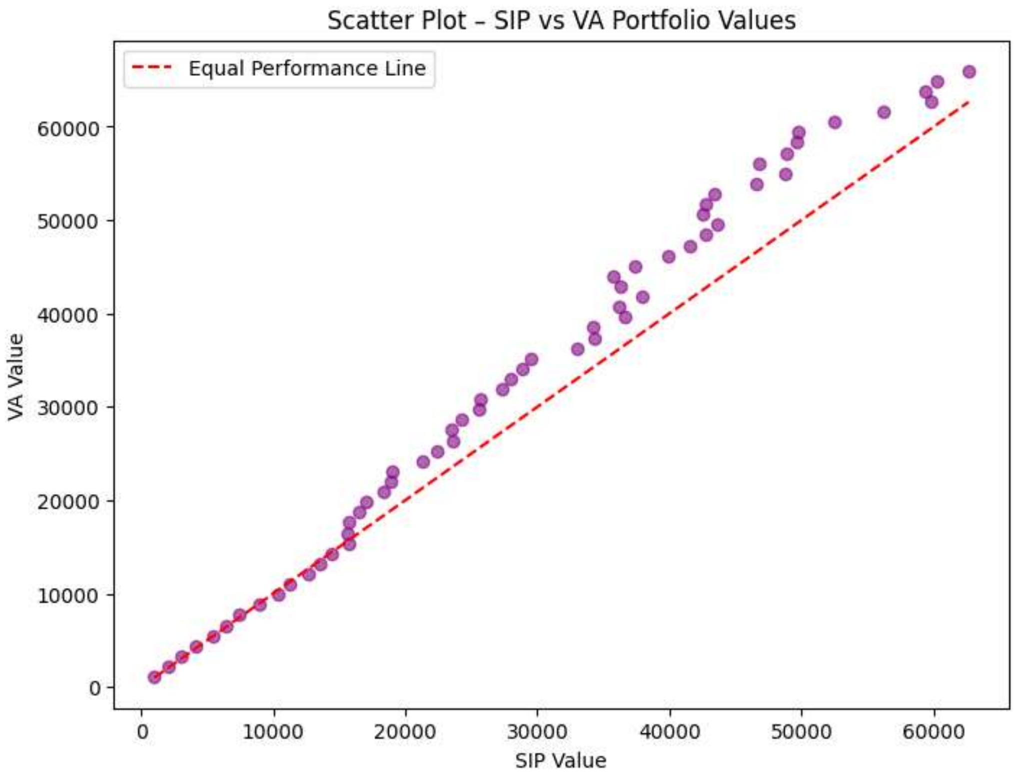

Figure 7: Scatter Plot – SIP vs VA Portfolio Values

Key Observation: Almost all points lie above the equal performance line, meaning VA portfolios outperform SIP portfolios in most months.

This study demonstrated that rule-based adaptive strategies such as Value Averaging (VA) and Systematic Investment Plan (SIP) consistently outperform static approaches like Lumpsum and Constant Share. Through the integration of a web-based intelligent system, the research compared portfolio outcomes under varying market conditions and applied regression forecasting for predictive analysis. The results confirmed that adaptive investment contributions improve long-term returns, reduce volatility, and provide superior risk-adjusted performance. The classifier and visualization modules further validated the strength of VA and SIP as dominant strategies, highlighting the role of AI in guiding retail investors toward evidence-based decision making.

Practical Implementation: The system can be deployed as a web-based investment advisory dashboard for retail users, enabling them to test different strategies before committing capital.

VI. Discussion

The findings of this research align with existing literature emphasizing the strengths of Value Averaging in volatile conditions and SIPs in stable markets. The results confirm that adaptive, rule-based models can provide superior performance under certain conditions. However, they also require higher monitoring and discipline

compared to SIPs.
The inclusion of regression forecasting, though basic, illustrates the potential of ML techniques in enhancing financial decision-making. Unlike high-cost proprietary robo-advisory platforms, this system provides a zero-cost, open implementation accessible to retail investors. The project contributes academically by demonstrating the integration of AI-inspired logic and modern web technologies for financial applications.

VII. Conclusion and Future Scope

The project successfully achieved its objectives by developing a working web-based investment simulator that integrates rule-based strategies, forecasting, and reporting. Comparative results confirmed that Value Averaging and its variants can outperform SIPs in volatile markets, while SIPs remain reliable in trending conditions. The system's accessibility, zero-cost implementation, and interactive interface highlight its value as a decision-support tool.

Future Scope:

1. Incorporating advanced ML models like LSTM, ARIMA, or neural networks for more accurate forecasting.
2. Adding sentiment analysis modules to correlate financial news or social media with stock movements.
3. Implementing diversification analysis using clustering and correlation heatmaps.
4. Deploying on cloud platforms like Azure for scalability and Excel integration.
5. Extending the simulator into a mobile application for broader accessibility.

References:

1. Edleson, M. E. (1991). Value Averaging: The Safe and Easy Strategy for Higher Investment Returns. Wiley.
2. Gupta, S., & Sharma, P. (2022). Comparative study of SIP vs Lumpsum in Indian Mutual Funds. International Journal of Financial Research, 13(4), 56–67.
3. Markowitz, H. (1952). Portfolio Selection. The Journal of Finance, 7(1), 77–91.
4. Patel, R., & Mehta, K. (2023). AI techniques in retail investment advisory systems. Journal of Intelligent Finance, 5(2), 112–128.
5. Aggarwal, R. (2023). Applications of regression models in stock price forecasting. Financial Data Science Review, 11(3), 89–101.

6. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Oversampling Technique. Journal of Artificial Intelligence Research, 16, 321–357.

7. Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. The Journal of Finance, 19(3), 425–442.

8. Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. The Journal of Finance, 48(1), 65–91.

9. Lo, A. W. (2004). The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective. Journal of Portfolio Management, 30(5), 15–29.

10. Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. Expert Systems with Applications, 42(1), 259–268.