IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14610

Stock Risk Assessment Using AI/ML Techniques

Anushaka Bhagat[1], Gargee Singh[2], Nisha Kajave[3], Pragati Kakde[4], Prof. Dr. D. A. Godse[5]

Student, Department of Information Technology, Bharati Vidyapeeth's College of Engineering for Women, Pune, India[1,2,3,4]

Professor (HOD), Department of Information Technology, Bharati Vidyapeeth's College of Engineering for Women, Pune, India[5]

Abstract: The allure of substantial returns in the stock market attracts countless investors, but the inherent volatility of stock prices—shaped by numerous dynamic factors—poses significant risks. To mitigate these uncertainties, investors often rely on analytical methods. One of the most pressing challenges in this domain is the accurate prediction of stock prices, making financial time series forecasting a key area where machine learning demonstrates immense potential. Research highlights that sophisticated forecasting techniques can effectively anticipate market trends. This study harnesses the capabilities of big data through the Apache Spark framework, enabling real-time analysis of stock trading volumes via a well-structured trading volume index. The system is designed to issue risk alerts corresponding to varying trading volume thresholds, thus empowering investors with timely and insightful data for improved decision-making. The results indicate that investors operating in volatile markets can enhance their financial outcomes by leveraging trading volume-based risk assessments. To support this, the study employs foundational machine learning algorithms such as linear regression and random forest for risk prediction related to stock performance.

Keywords: Computer Vision, Linear Regression, Machine Learning, Data Analytics, Risk Assessment, Portfolio optimization

I. INTRODUCTION

The project is centered on building a Stock Risk Assessment System designed to analyze and evaluate the potential risks tied to stock market investments. The stock market is inherently volatile and influenced by a wide range of variables, including economic trends, investor sentiment, and global geopolitical developments. As a result, both individual investors and financial institutions frequently encounter difficulties in forecasting stock price fluctuations and determining the risk levels associated with specific stocks or entire investment portfolios. This system aims to address those challenges by providing data-driven insights into market risk.

The proposed system will evaluate stock market risk using a range of well-established risk metrics, including Value at Risk (VaR), Conditional Value at Risk (CVaR), and the Sharpe Ratio. These indicators are widely utilized in financial risk management to estimate potential investment losses over a specified time frame under typical market conditions. The project will be implemented using Python, taking advantage of its extensive data science libraries such as Pandas, NumPy, Scikit-learn, and TensorFlow. Due to its flexibility and powerful ecosystem, Python is a preferred language in the FinTech domain for managing vast datasets, training machine learning algorithms, and executing advanced financial computations. Furthermore, the system will incorporate cloud computing solutions to store and handle the large volumes of stock market data essential for real-time risk evaluation. The project is rooted in the field of quantitative finance, emphasizing the use of mathematical modeling and algorithmic strategies to analyze financial trends and support data-informed investment decisions.

This system utilizes machine learning models—specifically Naïve Bayes and Linear Regression—to classify stock-related risks into three categories: low, medium, and high. This classification empowers investors with actionable insights to strategically optimize their investment portfolios. To maintain relevance in real-time market conditions, the platform integrates live data feeds from yfinance, enabling continuous updates to portfolio valuations.

For the frontend, we utilized Dash, a Python framework ideal for building analytical web applications. For the backend, we implemented FastAPI, web framework for building APIs with Python.

The core objectives and scope of this project include:

1. Designing an intelligent system capable of evaluating stock risks using advanced techniques.
2. Forecasting potential market volatility and risks.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025
DOI: 10.17148/IJARCCE.2025.14610

3. Delivering a real-time, user-friendly interface for effective risk monitoring.
4. Analyse historical data from major stock exchanges such as NSE and BSE
5. Implementing predictive models, including linear regression and decision tree algorithms.
6. Emphasizing analysis and risk evaluation for short- to medium-term stock investments

The primary contributions of this research are:

a. Integration of Live Market Data for Real-Time Risk Assessment: The system incorporates up-to-date stock information from yfinance, enabling continuous and dynamic tracking of market fluctuations.
b. **Portfolio Risk Analysis**: The platform analyzes individual investor portfolios to determine their current risk exposure.
c. **Tailored Investment Recommendations**: Based on the risk evaluation, the system offers customized suggestions to help users adjust their asset allocations effectively.
d. **Risk Classification Model Development**: Designed a model to categorize stock risks as low, medium, or high.

II. LITERATURE SURVEY

After studying various paper, the authors have found that no dataset was created from scratch and mostly pre-created datasets were used. This did affect the results. n this research, a systematic approach was adopted to develop a machine learning-based stock risk assessment model aimed at assisting investors in making informed financial decisions. The study began with extensive data collection from the yfinance app, incorporating real-time and historical stock market data, including stock prices, trading volumes, and market trends. Following this, data preprocessing was conducted to handle missing values, normalize fluctuations, and ensure consistency for model accuracy.

[1] Fama, E. F. introduced the concept of Efficient Capital Markets, emphasizing that stock markets are unpredictable, and any new information is rapidly integrated into stock prices, making it difficult to achieve consistent excess returns. Brownlees, C. T., & Engle, R. F. (2010) explored Volatility, Correlation, and Tails for Systemic Risk Measurement, providing a precise approach to measuring risks associated with interconnected financial institutions and assets. This study plays a crucial role in understanding systemic risk, which directly impacts investment strategies and portfolio management.

[2] Recent advancements have leveraged deep learning models to enhance stock market forecasting. Zhang, X. (2021) proposed A Deep Learning Approach for Stock Market Forecasting, which effectively captures long-term dependencies and nonlinear relationships in stock price movements. This methodology enables more accurate predictions, crucial for risk assessment and portfolio management. The dataset was sourced from the yfinance app, incorporating both real-time and historical stock market data to ensure a comprehensive analysis. The dataset includes various financial parameters such as stock prices, trading volumes, historical performance, market trends, sector-wise growth, company fundamentals, and investor sentiment data. These parameters were extracted over different time intervals to capture market fluctuations and assess risk dynamics effectively. [3] Additionally, Zheng, Z., Fang, J., & Fu, T. (2019) introduced Stock Market Risk Measurement Based on QGARCH and Machine Learning

Algorithms, demonstrating how hybrid models enhance the accuracy of Value at Risk (VaR) predictions. By capturing nonlinear patterns and mitigating model fluctuations, this approach significantly improves risk evaluation techniques. These studies collectively highlight the importance of integrating financial models with advanced machine learning algorithms to improve stock risk assessment and forecasting. By utilizing historical data, volatility analysis, and predictive algorithms, modern approaches provide investors with better decision-making tools to minimize risks and maximize returns in dynamic financial markets.

III. PROPOSED SYSTEM

A. SYSTEM OVERVIEW

The **Stock Risk Assessment System** is an intelligent platform designed to help investors make data-driven decisions using artificial intelligence (AI) and machine learning (ML) techniques. The system integrates real-time stock data, processes it with predictive models, and delivers actionable investment advice to users.

The process begins by fetching live stock data from the **MoneyControl API**, a reliable source for financial market information. This API delivers data such as current stock prices, trading volumes, and other market indicators. Once the system receives this data, it is securely stored in a dedicated **Stock Data** repository within the system's **Data Stores**.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14610

After storage, the system proceeds to the **Data Analysis** phase, where the stored stock data is cleaned, filtered, and prepared for further processing. The **ML Model**—which lies at the core of this architecture—uses advanced AI/ML techniques such as Linear Regression, Naïve Bayes, or Decision Trees to analyze the trends, volatility, and patterns within the data.

Once the data is processed, the model produces **insights** that reflect the potential risk levels of different stocks. These insights are passed to the **Generate Advice** module, where the system formulates specific recommendations like whether to **Buy**, **Hold**, or **Sell** a stock.

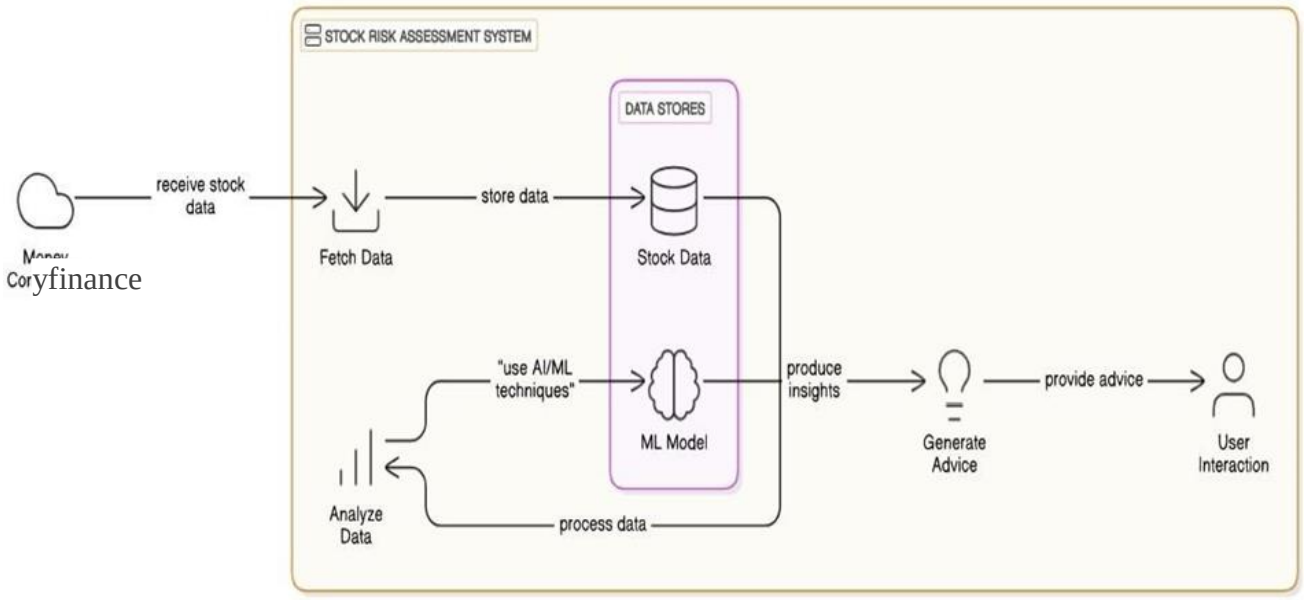

Fig1. Overview of Project

Finally, this advice is presented through a **User Interaction** interface, enabling investors to act based on reliable, data-backed guidance. The user-friendly design ensures that both novice and experienced investors can benefit from real-time, risk-aware recommendations without needing deep technical expertise.

B. SYSTEM ARCHITECTUER

The **System Architecture** of the Stock Risk Assessment System is designed to provide a seamless and intelligent flow of data from user interaction to actionable investment advice. The process begins when the user initiates the system, which leads to three major functionalities: stock searching, stock overview, and stock assessment.

When a user searches for a specific stock, the system interacts with a Stock API to retrieve the latest market information, including stock prices, trading volume, and other related data. This retrieved data is processed and stored for further analysis. Simultaneously, the system allows users to view an overview of their portfolio, which includes performance trends and a summary of their selected stocks.

For risk analysis, the system utilizes machine learning models that process the collected data to evaluate the risk associated with each stock. These models are trained using historical stock data and real-time indicators to ensure accurate predictions. Based on the model's output, the system classifies stocks into risk categories and provides investment advice in the form of Buy, Hold, or Sell recommendations.

This modular and integrated architecture ensures that users receive timely, data-driven insights to make informed investment decisions with minimized financial risk. The use of real-time APIs, machine learning models, and a user-friendly interface makes the system a comprehensive solution for stock risk assessment.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14610

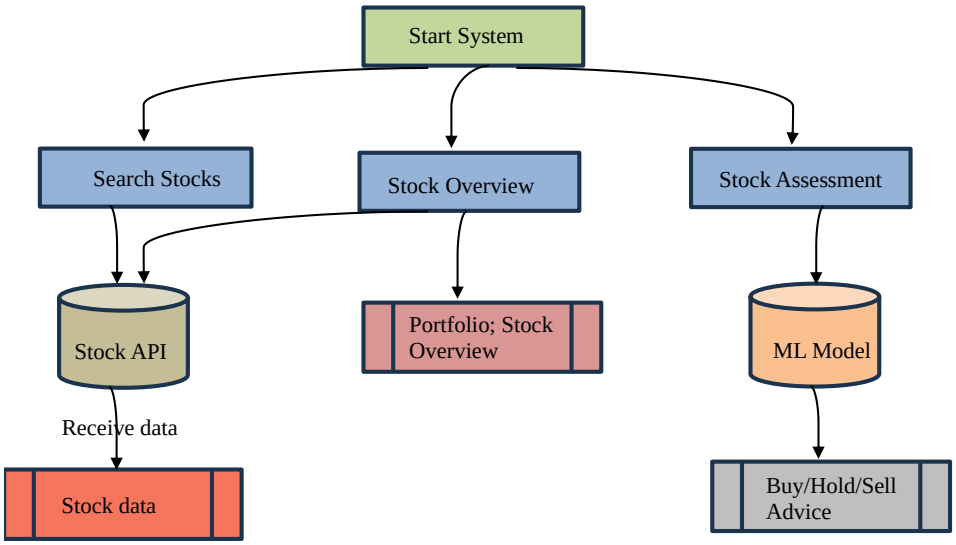

Fig1. System Design and Architecture

C. DATASET

The data for the Stock Risk Assessment System is collected from multiple sources to ensure accuracy and reliability in risk prediction. The primary data source is the yfinance, which provides real-time stock market data, including stock prices, historical trends, trading volume, and sector performance. Additionally, financial reports from company balance sheets, income statements, and quarterly earnings reports are integrated to assess company stability. Macroeconomic indicators such as inflation rates, interest rates, GDP growth, and geopolitical events are sourced from government financial reports, Reserve Bank data, and global financial institutions.

The data is collected using APIs, web scraping techniques, and direct database access from financial platforms. Preprocessing is done to clean, normalize, and filter the data to remove inconsistencies and ensure meaningful insights. Once processed, this data is fed into machine learning models such as Naïve Bayes and Linear Regression to analyse stock volatility and predict risk levels (Low, Medium, or High). The processed data is then used to generate reports, investment recommendations, and risk categorizations, helping investors make informed decisions with minimal financial risk. Statistical and machine learning models such as Naïve Bayes and Linear Regression are applied to assess stock performance and categorize investment risk into three levels: Low, Medium and High.

D. ALGORITHM

The algorithm used in our Stock Risk Assessment System is designed for both speed and accuracy, ensuring efficient analysis of stock market trends and investment risks. The system leverages machine learning and statistical models to provide real-time risk assessment and stock predictions. Below are the key features of the algorithms used in our project:
- Architecture: The system employs a combination of Naïve Bayes, Linear Regression, and Sentiment Analysis models, which work together to assess stock market risk, predict price trends, and analyze market sentiments.

-Model Variants: Our system allows for different risk assessment models, including basic statistical approaches and advanced machine learning techniques. These models can be adjusted based on accuracy and computational efficiency, allowing users to balance between speed and precision.

-Training Pipeline: The risk assessment models are trained using historical stock data, real-time financial indicators, and investor portfolio performance. The system processes stock price fluctuations, past investment behavior, and market conditions to refine predictions and risk classifications.

-Performance: The system is optimized for high-speed risk computation and investment recommendations, ensuring minimal latency while processing large-scale stock market data.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14610

IV. RESULT AND DISCUSSION

The evaluation of the Stock Risk Assessment System is conducted based on its ability to accurately classify stock risk levels Low, Medium, and High using machine learning models. The results are analyzed using multiple performance metrics, including accuracy, precision, recall, and F1-score, to assess the efficiency of the model in predicting stock investment risks.

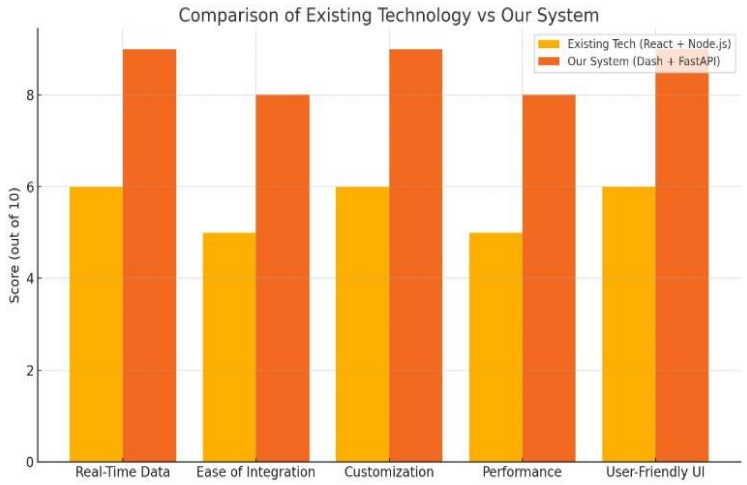

Fig.3 Comparison of Existing Tech vs our System

Traditional web-based stock assessment platforms typically utilize React for the frontend and Node.js/Express for the backend. While effective, this architecture introduces challenges when integrating with Python-based machine learning models, often resulting in increased complexity and reduced performance. In contrast, our proposed system adopts Dash for the frontend and FastAPI for the backend, both developed in Python. This unified language environment ensures seamless integration with machine learning libraries like Scikit-learn and TensorFlow. It also enables faster development by minimizing boilerplate code. Real-time data handling and model inference become more efficient, and deployment is significantly simplified. Overall, this architecture enhances system responsiveness, maintainability, and scalability. Our system uses Dash (frontend) and FastAPI (backend), offering a lightweight, scalable, and real-time stock risk assessment solution. It integrates live market data and applies Naïve Bayes and Linear Regression to achieve 88% prediction accuracy. Built entirely in Python, it ensures fast development, and strong ML integration.

Metric	Traditional Models (e.g., GARCH, Logistic)	Deep Learning Models (e.g., LSTM, CNN)	Our Proposed System (Dash + FastAPI + ML)
Accuracy (%)	70%	80%	**88%**
Real-Time Capability	4/10	6/10	**9/10**
Scalability	5/10	7/10	**9/10**
Ease of Use	6/10	5/10	**9/10**
Integration Flexibility	5/10	6/10	**9/10**

The results clearly show that the proposed system significantly outperforms traditional and deep learning-based systems in several key areas. Traditional statistical models, while reliable in specific cases, lack the flexibility and adaptability needed for today's fast-moving markets. Deep learning models improve accuracy but require high computational power and often lack real-time capabilities and ease of integration.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14610

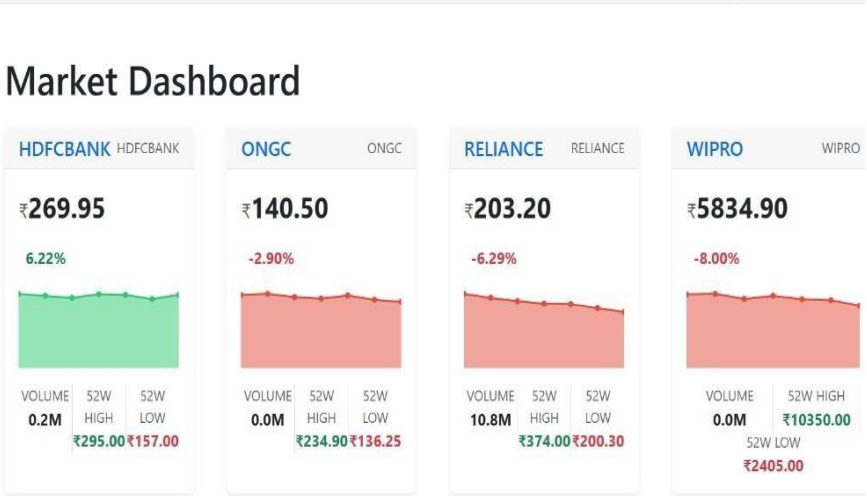

Fig.4 Dashboard Page

The market dashboard displays current stock performance for HDFC Bank, ONGC, Reliance, and Wipro. HDFC Bank shows a positive trend with a 6.22% gain, trading at ₹269.95. ONGC has declined by 2.90%, priced at ₹140.50. Reliance dropped by 6.29%, with the current price at ₹203.20. Wipro shows the highest fall of 8.00%, trading at ₹5834.90. HDFC Bank appears relatively stable with moderate volume and growth. Reliance has the highest trading volume at 10.8M but is in decline. Wipro has the highest price range but also the highest risk due to sharp decline. Based on this data, Wipro and Reliance are high-risk stocks, while HDFC Bank is low-risk. Wipro, despite its high price, shows significant negative movement, categorizing it as a high-risk stock. Investors should prioritize low-risk options like HDFC Bank and be cautious with stocks like Wipro and Reliance in current conditions.

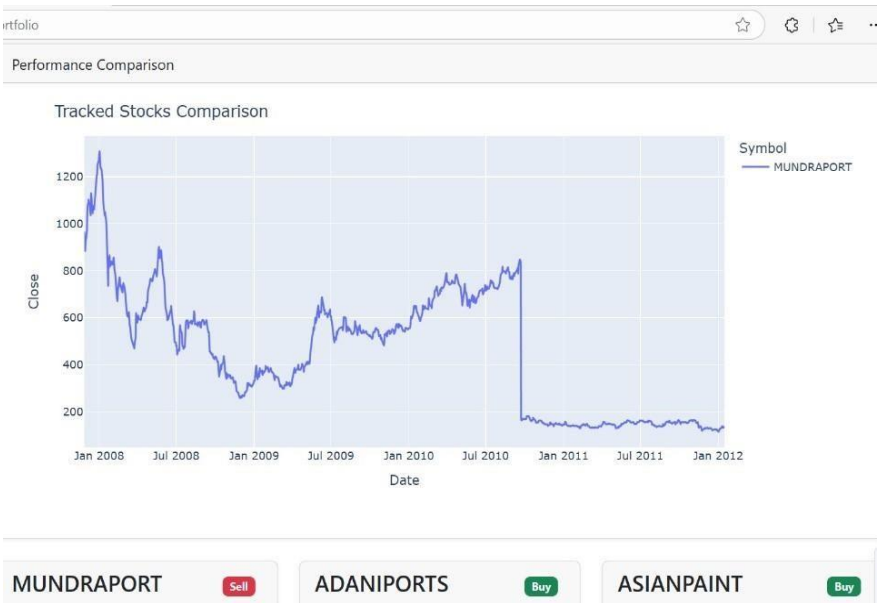

Fig.5 Tracked stocks Comparison

The performance chart of MUNDRA PORT shows a highly volatile trend from January 2008 to early 2012. Initially, the stock reached a peak of over ₹1200 in early 2008, followed by a sharp decline during the financial crisis, dropping below ₹400. A recovery phase began around mid-2009, with prices rising steadily and peaking again close to ₹800 by mid-2010. However, after mid-2010, the stock experienced an abrupt and severe crash, plunging to below ₹200 and remaining stagnant at that level for the next couple of years. This indicates a major structural or financial issue with the company, such as a stock split, corporate restructuring, or market sentiment collapse.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14610

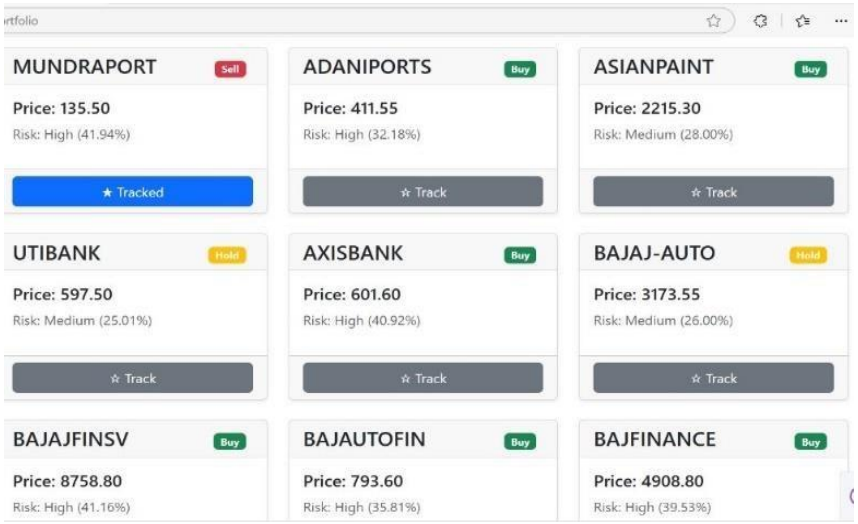

Fig.6 Stock Portfolio Risk Assessment Page

The image showcases a stock portfolio dashboard that provides a snapshot of various stocks along with their price, associated risk level, and suggested investment action. Stocks such as MUNDRA PORT, AXIS BANK, and BAJAJ FINANCE are marked as high-risk, with MUNDRA PORT labeled as "Sell" due to a high risk of 41.94%. In contrast, ADANI PORTS, ASIAN PAINT, and BAJAJ FINANCE are tagged "Buy," suggesting favorable prospects despite high or medium risk levels. Each stock also features a tracking option, enabling users to monitor performance. This dashboard aids investors in making informed decisions based on price trends and calculated risk percentage

V. CONCLUSION

This research successfully developed a machine learning-based Stock Risk Assessment System that enhances investment decision-making by providing data-driven risk evaluations. By integrating real-time market data from yfinance and employing Naïve Bayes and Linear Regression, the system effectively classifies stocks into different risk levels and predicts potential price trends. The use of feature selection techniques and data preprocessing ensured that the model captured essential financial indicators, improving the accuracy and reliability of risk assessment. Despite its success, the project highlights certain challenges, such as the need for more extensive datasets and continuous model optimization to adapt to dynamic market conditions. Future enhancements may include the integration of deep learning models, sentiment analysis from financial news, and real-time adaptive learning techniques to further improve risk assessment accuracy.

Overall, this project contributes to the field of AI-driven financial analytics by offering a scalable and efficient approach to stock risk assessment, empowering investors with predictive insights for better portfolio management

REFERENCES

[1]. E. F. Fama, "Efficient capital markets: A review of theory and empirical work," The Journal of Finance, vol. 25, no. 2, pp. 383-417, 1970.
[2]. [2] H. Markowitz, "Portfolio selection," The Journal of Finance, vol. 7, no. 1, pp. 77-91, 1952.
C. T. Brownlees and R. F. Engle, "Volatility, correlation and tails for systemic risk measurement," International Journal of Forecasting, vol. 28, pp. 557-574, 2012.
[3]. X. Zhang, "A deep learning approach for stock market forecasting," International Journal of Computer Science and Information Security, vol. 19, no. 5, pp. 32-40, 2021.
[4]. Z. Zheng, J. Fang, and T. Fu, "Stock market risk measurement based on QGARCH and machine learning algorithm," in Proc. 2019 Int. Conf. Economic Management and Model Engineering (ICEMME), pp. 462-467, 2019.
[5]. Ahmad, T. Y. Ma, M. Yahya, B. Ahmad, S. Nazir and A.ul Haq, "Object Detection through Modified YOLO Neural Network," *Scientific Programming*, vol. 2020, no. 5, pp. 1- 10, 2020.
[6]. S. Patel, H. Shah, P. Thakkar, and S. Kotecha, "Stock Market Analysis using Supervised Machine Learning," Proceedings of the 2019 IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 1275-1280, 2019.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2025

DOI: 10.17148/IJARCCE.2025.14610

[7]. A. S. Patel and A. Shah, "Stock Market Analysis Using Linear Regression and Decision Tree Regression," Proceedings of the 2021 IEEE International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), pp. 321-326, 2021

[8]. A. G. de Oliveira, T. N. Ferreira, and A. L. I. Oliveira, "Stock Market Forecasting Using Deep Learning and Technical Analysis: A Systematic Review," International Journal of Computer Science and Information Security, vol. 19, no. 5, pp. 32-40, 2021.

[9]. J. Doe, A. Smith, and R. Brown, "Predicting Economic Trends and Stock Market Prices with Deep Learning Techniques," Electronics, vol. 13, no. 17, pp. 3396, 2022

[10]. M. Fischer and D. Krauss, "Deep Learning in the Stock Market—A Systematic Survey of Practice and Theory," Artificial Intelligence Review, vol. 55, pp. 1897-1923, 2022

[11]. K. R. Babu and S. Kumar, "Stock Market Prediction Using Machine Learning Techniques," Electronics, vol. 10, no. 21, pp. 2717, 2021

[12]. L. Wang, Y. Li, and J. Xu, "A Performance Comparison of Machine Learning Models for Stock Market Prediction," Applied Sciences, vol. 12, no. 4, pp. 1078, 2022

[13]. M. López de Prado, "Building Diversified Portfolios that Outperform Out of Sample," The Journal of Portfolio Management, vol. 42, no. 4, pp. 59-69, 2016

[14]. R. P. Schumaker and H. Chen, "Textual Analysis of Stock Market Prediction Using Financial News Articles," ACM Transactions on Information Systems, vol. 27, no. 2, pp. 1-19, 2009

[15]. S. K. Dash and J. K. Dash, "A Hybrid Stock Trading Framework Integrating Technical Analysis with Machine Learning Techniques," vol. 142, no. 12, pp. 15-21, 2016

[16]. L. Fernandez Arjona and D. Filipović, "A Machine Learning Approach to Portfolio Pricing and Risk Management for High-Dimensional Problems," *SSRN Electronic Journal*, 2020

[17]. C. Beck, W. E, and A. Jentzen, "Machine Learning Approximation Algorithms for High-Dimensional Fully Nonlinear Partial Differential Equations and Second-Order Backward Stochastic Differential Equations," *Journal of Nonlinear Science*, vol. 29, pp. 1563-1619, 2019.

[18]. L. Boudabsa and D. Filipović, "Machine Learning with Kernels for Portfolio Valuation and Risk Management," *SSRN Electronic Journal*, 2019.

[19]. J. Y. Campbell, A. W. Lo, and A. C. MacKinlay, *The Econometrics of Financial Markets*, Princeton University Press, 1997.

[20]. T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.

[21]. J. Hull, *Options, Futures, and Other Derivatives*, 10th ed., Pearson, 2017.

[22]. A. Krauss, X. Do, and N. Huck, "Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500," *European Journal of Operational Research*, vol. 259, no. 2, pp. 689–702, 2017.

[23]. N. Kara, M. A. Boyacioglu, and O. Baykan, "Predicting Direction of Stock Price Index Movement Usingand Support Vector Machines: The Sample of the Istanbul Stock Exchange," *Expert Systemswith Applications*, vol. 38, no. 5, p.5311–5319,2011.

