MULTIDISCIPLINARY RESEARCH INSTITUTE INDIA

International Journal on Advanced Computer Theory and Engineering

ISSN: 2319 - 2526

Volume 15 Issue 2s (2026) | Pages: 152-162

AI Based Equity Portfolio Management System

Prahlad Shanbhag[1], Sandeep Borkar[2], Kunal Gangurde[3], Suchita Ghute[4]

[1,2,3] Student, Department of Artificial Intelligence and Data Science, GSMCOE Balewadi, Maharashtra, India

[4]Professor, Department of Artificial Intelligence and Data Science , GSMCOE Balewadi, Maharashtra, India

Peer Review Information

Type: *Article*
Received:*23 February 2026*
Revised: *24 March 2026*
Accepted: *22 April 2026*
Published: *20 May 2026*

Abstract

Portfolio optimization is a fundamental challenge in financial decision-making, especially in dynamic and volatile markets such as the Indian stock market. Investors often struggle to construct portfolios that balance risk and return effectively due to fluctuating market conditions, sector-specific variations, and unpredictable economic factors. This study presents a structured and data-driven approach for designing optimized equity portfolios across thirteen major sectors listed on the National Stock Exchange (NSE) of India. Using historical stock price data from 2017 to 2021, the research explores three widely recognized portfolio construction methodologies: the Equal Weight Portfolio (EWP), the Minimum Risk Portfolio (MRP), and the Mean- Variance or Optimum Risk Portfolio (ORP). Each method employs mathematical and statistical techniques to evaluate daily returns, annual volatility, covariance, and correlation matrices, enabling a comprehensive understanding of risk- return dynamics. The study incorporates the top ten stocks from each sector based on free-float market capitalization, making the analysis both sector-specific and representative of real market conditions. These portfolios are then evaluated on actual stock performance data from the year 2022 to determine their practical effectiveness under real- world market fluctuations. Such a two-phase evaluation—training on historical data and testing on real unseen data.

Keywords: Portfolio Optimization; Equity Portfolio Management; Risk-Return Analysis; Mean-Variance Optimization; Minimum Risk Portfolio; National Stock Exchange (NSE); Financial Decision-Making; Stock Market Analysis; Portfolio Risk Management; Investment Strategies

How to Cite This Article

Shanbhag, P., Borkar, S., Gangurde, K., & Ghute, S. (2026). *AI Based Equity Portfolio Management System*. *International Journal on Advanced Computer Theory and Engineering,* 15(2s), 152–162.

Introduction

Portfolio optimization is the process of identification of a set of capital assets and their respective weights allocation such that the risk-return pairs are optimized. The optimization problem is further confounded due to the requirement of the estimation of the future returns of stocks that exhibit volatility in their prices.

This paper presents a methodical approach to building robust stock portfolios in thirteen key sectors of the Indian economy. For each of these thirteen sectors, the top ten stocks based on their free-float market capitalization in the national stock exchange (NSE) are identified (NSE Website). The ticker names of these 130 stocks are used to automatically scrape their historical prices from the Yahoo Finance website (Yahoo Finance Website). Using these historical prices over five years, three different portfolios are designed following the equal weight allocation, minimum risk portfolio allocation, and mean-variance portfolio allocation.

Literature Survey

Portfolio optimization has been one of the most widely researched areas in finance, operations research, and computational economics for several decades. The earliest and most influential contribution to this field was made by Harry Markowitz in 1952, who introduced the Modern Portfolio Theory (MPT). His work demonstrated that the optimal selection of assets should be based on the trade-off between expected return and risk, quantified through variance and covariance. Since then, the literature has evolved significantly with researchers exploring various statistical, econometric, machine learning, and deep learning techniques to overcome the limitations of classical models. Traditional econometric models such as ARIMA, VAR, and multivariate regression have been extensively used to model stock prices and forecast future returns. These models rely heavily on time-series stability and linear relationships between variables. However, the real- world financial market shows non-linearity, irregular patterns, and volatility clustering, making it difficult for classical models to consistently achieve high prediction accuracy.

underlying to address these limitations, recent studies have increasingly incorporated machine learning techniques, including Random Forest, Support Vector Regression (SVR), Gradient Boosting, and ensemble learning methods. These models are more flexible and can capture non-linear relationships. Still, as financial markets involve complex dependencies, machine learning alone sometimes fails to capture long-term temporal structures and sudden shocks. This led to the broader adoption of deep learning approaches, especially LSTM and CNN-LSTM hybrid models, which have shown promising results for stock price prediction. LSTM networks, with their ability to retain long-term dependencies, are particularly effective in identifying patterns in noisy historical data. Many researchers (such as Mehtab S Sen, 2021; Chatterjee et al., 2021) demonstrate that deep learning-based forecasting tends to outperform traditional econometric models, especially for volatile markets. Beyond prediction models, literature has explored several optimization- based portfolio construction techniques. Apart from mean-variance optimization, researchers have proposed risk- parity portfolios, hierarchical risk parity (HRP), eigen portfolios using PCA, genetic algorithms, fuzzy logic-based models, swarm intelligence algorithms, and reinforcement learning frameworks. These methods attempt to optimize the asset allocation process by either improving the estimation of inputs (returns, risk, covariance) or incorporating dynamic adjustment rules that respond to market fluctuations. A significant portion of literature focuses on applying these portfolio optimization techniques to specific financial markets, such as emerging markets. Studies show that emerging markets, like India, often behave differently from developed markets due to liquidity variations, sector imbalances, and market sentiments influenced by macroeconomic conditions. Therefore, sector-wise portfolio optimization—analyzed in this study—has been recognized as an important approach to understanding risk-return characteristics more precisely.

Methodology

As discussed in Section 1, thirteen sectors are selected from the NSE, India. The sectors are auto, banking, consumer durables, financial services, fast-moving consumer goods (FMCG), information technology (IT), media, metal, oil S gas, pharma, public sector banks, private banks, and realty. Based on the free-float market capitalization in the NSE, the top ten stocks for each of the above mentioned sectors are identified. The sector-specific portfolios are built using those stocks. In the next step, to get an understanding of the association strength between the close prices of a pair of stocks in a sector, their covariance and correlation matrices are calculated using the training dataset. A strong association of a pair is indicated by a high covariance value between them and vice-versa. The matrices are computed using the cov and corr Python functions. A good portfolio aims to optimize the return while minimizing its risk. To minimize the risk, the identification of stocks with low correlation among them is required. This helps in achieving a higher diversity in the portfolio, thereby reducing its risk.

Advanced Data Preprocessing Framework

For each sector, the prices of the stocks are acquired from the Yahoo Finance site using the Yahoo Financials function. The parameters passed in the function are the ticker name, and the date range from Jan 1, 2017, to Dec 31, 2022, with "daily" frequency. The extracted stock data have the following attributes:

- open
- high
- close
- volume
- adjusted close

This work is based on univariate analysis and hence the variable 'Close' is chosen as the variable of interest, while the remaining variables are ignored. The close values of the ten stocks, from Jan 1, 2017, to Dec 31, 2021, are used for building the portfolios. The portfolios are tested for their returns from Jan 1, 2022, to Dec 30, 2022.

Algorithm Selection Strategy

The algorithm selection strategy in the AI-based Equity Portfolio System is designed to analyze financial data characteristics and automatically determine suitable modeling approaches for prediction, risk assessment, and portfolio optimization tasks. Based on the nature of the target variable and problem objective, the system dynamically selects appropriate machine learning algorithms. Additionally, the system can be extended to include **portfolio-level visualizations**, such as asset allocation charts, risk-return trade-off graphs, and performance trends over time, enabling users to evaluate overall portfolio effectiveness.

This led to the broader adoption of deep learning approaches, especially LSTM and CNN-LSTM hybrid models, which have shown promising results for stock price prediction. LSTM networks, with their ability to retain long- term dependencies, are particularly effective in identifying patterns in noisy historical data. Many researchers (such as Mehtab S Sen, 2021; Chatterjee et al., 2021) demonstrate that deep learning-based forecasting tends to outperform traditional econometric models, especially for volatile markets.Following are the top ten stocks of this sector based on their free-float market capitalization in the NSE, as per the report released on Dec 30, 2022. The figures below represent the contributions in percentage, of the stocks (computed based on the market capitalization) to the overall index of the banking sector. The ticker name is also mentioned with a pair of parentheses just beside the name of the respective stock.

HDFC Bank (HDFCBANK): 28.66 ICICI Bank (ICICIBANK):23.54

Kotak Mahindra Bank (KOTAKBANK): 10.18 Axis Bank (AXISBANK): 10.01

State Bank of India (SBIN): 9.85 IndusInd Bank (INDUSINDBK): 5.91 Bank of Baroda (BANKBARODA):

2.62 AU Small Finance Bank (AUBANK): 2.49 Federal Bank (FEDERALBANK): 2.38 IDFC First Bank (IDFCFIRSTB): 1.55

To address these limitations, recent studies have increasingly incorporated machine learning techniques, including Random Forest, Support Vector Regression (SVR), Gradient Boosting, and ensemble learning methods. These models are more flexible and can capture non-linear relationships. Still, as financial markets involve complex dependencies, machine learning alone sometimes fails to capture long-term temporal structures and sudden shocks. For price prediction and return forecasting involving continuous target variables, the system employs regression- based models, including:

- Linear Regression (baseline model for trend estimation)
- Decision Tree Regressor (captures non-linear relationships in stock data)
- Random Forest Regressor (ensemble learning for improved accuracy and robustness)
- Support Vector Regression (effective for handling high- dimensional financial data)

For **regression** tasks with continuous targets, available algorithms include:

- Linear Regression (ordinary least squarePredicted vs Actual Scatter Plot – comparison with a reference diagonal line (Fig - 4)
- Residual Plot – visualization of error patterns to detect bias and variance.
- Residual Distribution Histogram – analysis of prediction error distribution

- Error Metrics Visualization – comparison of MSE, RMSE, MAE, and RZ scores

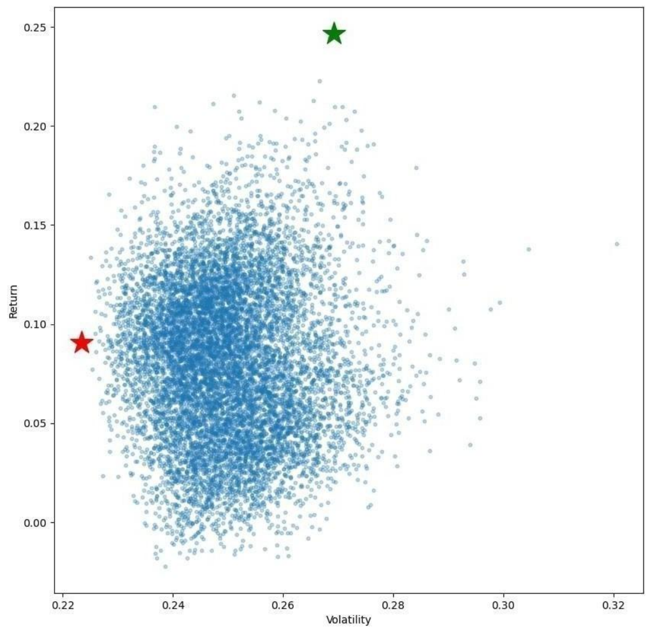

Fig. 1. *The efficient frontier of Auto sector.*

Training Process

The model training process executes selected algorithms using either default configurations for rapid analysis or optimized hyperparameters for improved performance. Hyperparameter tuning is performed using techniques such as Grid Search with cross-validation to ensure robust model selection. The system provides real-time progress updates to the user during model execution, enabling transparency in processing and improved user interaction.

Upon completion, the system evaluates model performance using appropriate financial and statistical metrics. For prediction tasks, metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and RZ score are computed, while classification-based decisions (e.g., buy/sell/hold signals) are assessed using accuracy, precision, recall, and F1-score. These results are presented through interactive visualizations to facilitate better interpretation of model outcomes.

Model Performance Visualization

The AI-based Equity Portfolio System provides comprehensive visualization capabilities to enable users to intuitively interpret model performance and make informed investment decisions. The visualization module utilizes interactive web-based charts to present analytical results, facilitating better understanding of predictions, risk, and portfolio behavior . intuitively interpret model performance and identify areas for improvement. For **price prediction and return forecasting**, the system generates visualizations such as predicted versus actual value plots, which illustrate the alignment between model predictions and real market data. A strong correlation between predicted and actual values indicates the model's capability to capture underlying market trends, while deviations highlight potential areas for improvement due to volatility or data limitations.

To further assess model reliability, **residual analysis** is performed. Residual plots and error distribution histograms help identify patterns such as bias, variance, or outliers in predictions. Additionally, key evaluation metrics including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and RZ score are computed and visualized to provide a quantitative measure of model performance.

The visualization module leverages Chart.js to generate interactive, web-based plots that facilitate model evaluation and comparisonThe risk and return values for the three portfolios, computed using stock prices over the training period, are depicted in Table 8. It is evident that the equal weight portfolio has yielded the lowest return and the minimum risk portfolio has exhibited the least risk. The optimum risk portfolio has produced the highest return, while the equal weight portfolio has exhibited the highest risk.

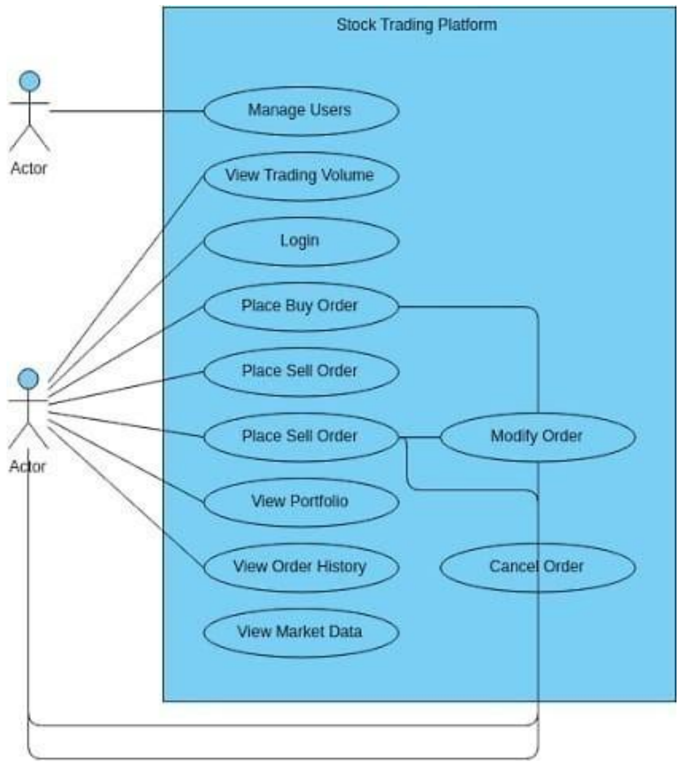

Fig. 2. *Unified Modeling Language.*

For **classification-based investment decisions** (e.g., buy, sell, or hold signals), the system utilizes visualization tools such as confusion matrices, ROC curves, and precision-recall curves. These visualizations enable users to assess classification accuracy, trade-offs between true and false positives, and performance under imbalanced market conditions. Detailed classification reports further provide insights into precision, recall, and F1-score for each decision category.

In addition to model-level evaluation, the system supports **portfolio-level visualizations**, including asset allocation charts, cumulative return graphs, and risk-return trade-off plots. These visual tools help users understand portfolio behavior over time and evaluate the effectiveness of investment strategies.

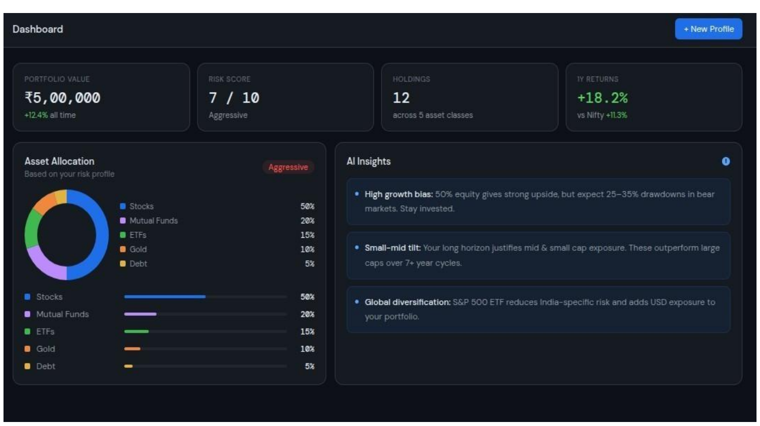

Fig. 3. *Dashboard*

System Architecture

The AI-based Equity Portfolio System is designed as an intelligent decision-support platform that guides users through the complete investment analysis and portfolio management lifecycle in an intuitive manner. The system follows a three-tier architecture consisting of a frontend interface for user interaction, a backend processing layer for data analysis and model execution, and a database layer for managing financial data and user information. This architecture ensures scalability, responsiveness, and efficient handling of real-time user requests.

Data Ingestion and Validation

The workflow begins with a structured data ingestion process, where users upload datasets in CSV or JSON format through an interactive interface. Upon upload, the system performs validation checks to ensure data consistency and usability.

These checks include verification of file structure, consistency of columns, and adherence to size constraints. Each dataset is assigned a unique identifier and stored systematically, allowing users to maintain different versions of their datasets. This approach supports traceability and reproducibility of experiments.

Intelligent Pipeline Configuration

DataSage incorporates a guided mechanism for selecting target variables and configuring preprocessing steps. The system analyses dataset characteristics to recommend suitable target columns based on data type and distribution.

Users can apply preprocessing techniques such as handling missing values, encoding categorical variables, and managing outliers through a unified interface. The system also allows intermediate states of processed datasets to be saved, enabling users to revisit . different preprocessing configurations. This enhances flexibility and supports iterative experimentation.

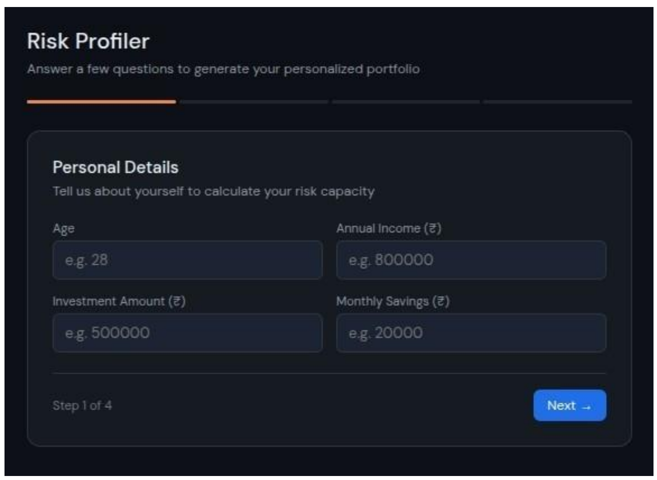

Fig. 4. *Risk Profiler*

Automated Model Training and Validation

Based on the selected target variable, the system automatically determines whether the task is classification or regression and suggests appropriate algorithms accordingly. Multiple models can be trained in parallel to improve efficiency.

To ensure reliable performance, the system incorporates cross-validation and hyperparameter tuning techniques. These methods help in selecting models that generalize well to unseen data while minimizing overfitting. Users are provided with real-time feedback during the training process, improving usability and engagement.

Visualization and Model Export

The final stage focuses on model evaluation and interpretability. The platform generates visual outputs such as confusion matrices, ROC curves, precision-recall curves, and residual plots, depending on the problem type.

Testing And Validation

Unit Testing of Core Modules

Individual components of the system were tested to verify their functional correctness under various conditions.

The preprocessing modules were evaluated using synthetic datasets containing edge cases such as missing The data preprocessing module was tested to validate operations such as data cleaning, handling missing values, normalization, and feature engineering. Test cases ensured that the system correctly processes raw financial data and prepares it for model training without introducing inconsistencies.

values, outliers, and mixed data types. These tests ensured that preprocessing operations produced consistent and meaningful outputs without introducing bias or unintended data leakage.

Additionally, the semantic type detection mechanism was validated using a diverse set of input columns, including identifiers, timestamps, and categorical text data. The system demonstrated high accuracy in correctly identifying column types, supporting appropriate preprocessing and model selection.

The dataset health scoring mechanism was also assessed by comparing automated scores with manually evaluated data quality indicators. This helped ensure that the generated scores accurately reflect the suitability of datasets for machine learning tasks.

Integration and System Testing

The AI-based Equity Portfolio System follows a multi-tier architecture; therefore, integration testing was conducted to ensure seamless interaction between all system components.

Integration with **financial data sources and APIs** was validated to ensure accurate and consistent retrieval of stock prices, historical data, and market indicators. The system's ability to process and update data in near real- time was also tested to support timely investment decision.

The communication between the frontend interface and backend services was thoroughly evaluated to verify correct data flow during key operations such as market data retrieval, portfolio creation, model execution, and prediction generation. The system was tested to ensure it can handle concurrent user requests efficiently while maintaining responsiveness, especially during real-time data processing and analysis

Model Validation Strategy

The machine learning pipeline incorporates standard validation techniques to ensure reliable model performance. Cross-validation methods were used to evaluate model generalization and reduce the risk of overfitting. During model training, hyperparameter tuning techniques were applied to identify optimal configurations for different algorithms.

Performance metrics such as accuracy, precision, recall, and F1-score for classification tasks, along with mean squared error, root mean squared error, and RZ score for regression tasks, were computed and verified for correctness. These metrics provide a comprehensive assessment of model effectiveness.

User Interface Evaluation

The user interface was evaluated to ensure usability, accessibility, and responsiveness across different environments.

The system was tested on multiple devices and screen sizes to verify consistent performance of visual components such as charts and evaluation dashboards.

Data upload functionality was also assessed to ensure smooth handling of different file formats and sizes.

Overall, the interface was designed to maintain clarity and ease of use, enabling users with minimal technical expertise to navigate the platform effectively.

Results And Evaluation

Benchmark Dataset Performance

The daily return, which is the percentage change in consecutive daily close prices, for each stock is computed on the training data using the pct_change function of Python. The annual return for each stock is computed as the weighted sum of the daily returns for the stock. Using the daily return values, the daily and annual volatility of the stocks are then calculated, using (1) and (2).

$Daily\ volatility$ (v) = $D\ standard\ deviation\ of\ daily\ returns$ (1)

$Annual\ volatility\ (v)$= D√250 , assuming 250 working days/year for stock market, (2) Since annual volatility indicates price variability, it provides investors with a measure of the risk associated with a stock. The std function of Python function is used to calculate the volatility of stocks.

Table 1. The portfolio weights for the auto sector stock

Stock	**EWP**	**MRP**
MSM	0.1	0.064107
MARUTI	0.1	0.163289
TATAMOTORS	0.1	0.028795
EICHERMOT	0.1	0.122902
BAJAJ-AUTO	0.1	0.242138
HEROMOTOCO	0.1	0.072284
TIINDIA	0.1	0.155767
TVSMOTOR	0.1	0.149395
BHARATFORG	0.1	0.000088
ASHOKLEY	0.1	0.001234

BAJAJ-AUTO receives the highest allocation as per the minimum risk portfolio while TIINDIA receives the highest allocation as per the optimum risk portfolio. Figure 3 depicts the weight allocation to the auto sector stocks done by the optimum risk portfolio.

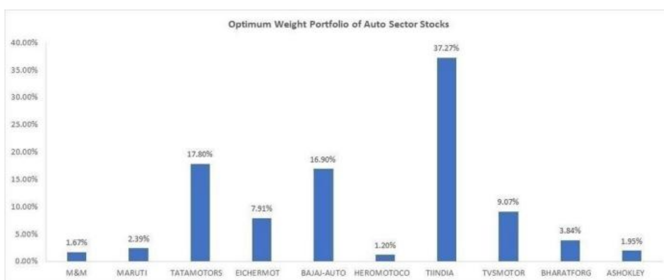

Fig. 3. *The ORP Portfolio Weights for banking sector.*

Table 2. The return and risk of the auto sector portfolios

Metric	**EWP**	**MRP**	**MVP/ ORP**
Portfolio annual return(%)	8.147	8.77	27.94
Portfolio annual risk (%)	24.46	22.45	25.60

The risk and return values for the three portfolios, computed using stock prices over the training period, are depicted in Table 3 above. It is evident that the equal weights portfolio has yielded the lowest return and the minimum risk portfolio has exhibited the least risk.

The optimum risk portfolio has produced the highest return and the highest risk. It should be noted the time required to complete key stages, including data preprocessing, feature engineering, model selection, and evaluation.

As shown in Table 2, the automated approach significantly reduces the time required at each stage of the pipeline. Tasks such as data cleaning and feature encoding, which typically require substantial manual effort, are completed within seconds or minutes.

The total pipeline execution time was reduced from approximately several hours in a manual workflow to a few minutes using the platform, representing a substantial improvement in efficiency. This reduction enables non-expert users to quickly build and evaluate machine learning models without extensive programming knowledge.

The AI-based Equity Portfolio System adopts a comprehensive model validation strategy to ensure reliability, robustness, and generalization of predictive models in dynamic financial markets. Given the time-dependent nature of stock market data, the system emphasizes validation techniques that preserve temporal order and prevent data leakage.

For price prediction and return forecasting, the system utilizes time-series–aware validation methods, such as train-test splits based on chronological order and rolling or expanding window validation. This approach ensures that models are always evaluated on future data, closely simulating real-world deployment scenarios.

To enhance model robustness, k-fold cross-validation is applied where appropriate, particularly for non-sequential datasets or feature-engineered inputs. This helps in reducing variance and ensures that the model performance is not dependent on a specific data split.

The system evaluates models using a combination of statistical and financial performance metrics. Regression models are assessed using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and RZ score, while classification models (e.g., trading signals) are validated using accuracy, precision, recall, and F1-score. In addition, financial metrics such as cumulative returns and risk-adjusted measures can be considered to evaluate practical investment performance.

To prevent overfitting, the system incorporates techniques such as cross-validation, regularization, and controlled hyperparameter tuning. Model performance is compared across multiple algorithms, and the best-performing model is selected based on validation results.

The system also considers risk-adjusted validation metrics, such as Sharpe ratio or volatility measures, to evaluate model performance from an investment perspective rather than purely statistical accuracy. This ensures alignment with real-world financial objectives.

Additionally, sensitivity analysis is performed to understand how changes in input features (such as stock prices, volume, or external indicators) affect model predictions. This helps in identifying critical factors influencing decision-making and ensures model reliability under varying conditions.

Finally, periodic model revalidation and retraining is incorporated to adapt to changing market conditions. As financial markets are highly dynamic, continuous monitoring and updating of models are essential to maintain predictive performance over time.

Model Export and Reliability

The platform also ensures reliability and reproducibility of trained models. All generated models were successfully exported in serialized format, including both the trained estimator and the associated preprocessing pipeline.

Validation tests confirmed that the exported models produced consistent predictions when deployed in external environments. This demonstrates that the system supports seamless transition from model development to deployment, reinforcing the practical applicability of the platform.

Research Gaps And Future Scope

While the AI-based Equity Portfolio System addresses several challenges in intelligent investment decision-making, our analysis identifies remaining gaps that present opportunities for further enhancement and future research contributions.

Current Limitations

The AI-based Equity Portfolio System provides comprehensive visualization capabilities to enable users to intuitively interpret model performance and make informed investment decisions. The visualization module utilizes interactive web-based charts to present analytical results, facilitating better understanding of predictions, risk, and portfolio behavior. Tasks (classification and regression), leaving unsupervised learning techniques such as clustering, dimensionality reduction, and anomaly detection unavailable to users. Deep learning capabilities are not included, limiting the platform to traditional machine learning algorithms unable to process complex data types like images, text, or time series that benefit from neural network architectures. Model interpretability features provide basic performance metrics but lack advanced explainability techniques such as SHAP values, LIME explanations that help users understand model decision- making processes.

Identified Research Gaps

Literature analysis in the domain of AI-driven equity portfolio management reveals several underexplored areas. While numerous platforms provide algorithmic trading and portfolio optimization capabilities, explainability and transparency remain limited, with most systems failing to justify why specific assets are selected or why certain allocation strategies are recommended. This lack of interpretability reduces user trust, especially among retail investors and non-expert users.Real-time adaptability and learning is also insufficiently addressed. Many systems operate on static or periodically updated models, lacking the ability to continuously learn from streaming market data, news sentiment, or macroeconomic indicators. This limits their responsiveness to rapidly changing market conditions.

Furthermore, integration of alternative data sources such as social media sentiment, financial news, and global economic indicators remains limited. Most platforms primarily depend on historical price data, which restricts the predictive power of AI models in capturing market.

Future Enhancement Directions

Based on the identified research gaps, several enhancement directions are proposed for the AI-based Equity Portfolio System. Expanding analytical capabilities to incorporate advanced machine learning models, including deep learning architectures and time-series forecasting techniques, would significantly improve predictive accuracy for stock price movements and portfolio performance. Integration of hybrid models combining statistical methods with AI approaches can further enhance robustness in volatile market conditions.

To address the lack of transparency, the implementation of explainable AI (XAI) techniques such as SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model- agnostic Explanations), and counterfactual explanations is recommended. These methods would enable users to understand the rationale behind asset selection, risk assessments, and allocation strategies, thereby improving trust and usability.transfer learning capabilities, and time series forecasting methods would broaden applicability across domains. Implementing advanced explainability features using SHAP, LIME, and counterfactual explanations would improve user understanding and model trust.

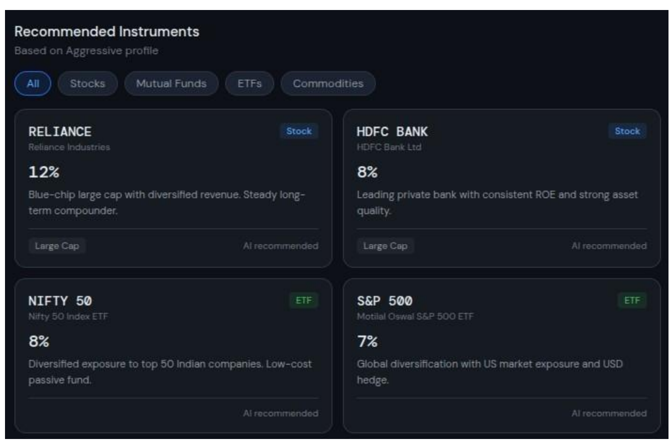

Fig. 5. *Recommendation Instruments*

Another key limitation is the lack of **personalized investment strategies**. Most platforms rely on static risk profiles and generalized models, failing to incorporate individual investor preferences, behavioral patterns, financial goals, and dynamic risk tolerance. This restricts the effectiveness of portfolio recommendations in real- world scenarios.

The integration of **real-time and alternative data sources** is also insufficient. While traditional models depend heavily on historical price data, limited attention is given to incorporating financial news, social media sentiment, and macroeconomic indicators, which play a crucial role in influencing market behavior.

Furthermore, existing systems often lack **adaptive learning capabilities**, operating on static or periodically updated models rather than continuously learning from evolving market conditions. This reduces their ability to respond effectively to sudden market fluctuations and emerging trends.

Conclusion

Designing a portfolio for optimizing the return and risk is a very challenging task. In this paper, two approaches to portfolio design are presented. These approaches are the equal-weight portfolio design and the mean-variance portfolio (also known as the optimum risk portfolio) design. Thirteen important sectors listed on the NSE of India are chosen, and the top ten stocks from each sector are identified based on free-float market capitalization. Using the historical stock prices for each stock from Jan 1, 2017, to Dec 31, 2021, two portfolios are designed for each sector. The portfolios are tested over the stock price data from Jan 1, 2022, to Dec 31, 2022. The evaluation of the portfolios is done based on the annual returns yielded by them. It is observed that the returns of the equal-weight portfolio are higher for seven sectors. The sectors for which the equal-weight portfolios have yielded higher returns are banking, consumer durables, financial services, media, pharma, public sector banks, and private banks. However, the optimum risk portfolios produced higher returns than their equal-weight counterparts for six sectors. These sectors are auto, FMCG, IT, metal, oil S gas, and realty. There are five sectors for which both portfolios are found to have produced negative returns over the year 2022. These sectors are consumer durables, IT, media, pharma, and realty. Investors should avoid investing in the stocks of these sectors to avoid loss till the sentiments in the Indian stock market takes a positive turn. Overall, the public sector banks sector yielded the highest return and is found to be the most profitable one. On the other hand, the IT sector has yielded the highest negative return, and hence, this sector looks to be very risky and should be avoided by investors until the Indian stock market sentiments enter into a bull phase. Comparing the performances of the portfolios designed based on other approaches such as hierarchical risk parity (HRP), Eigen portfolios based on principal component analysis, Black- Litterman portfolio, and hierarchical equal risk contribution (HERC) portfolios on these thirteen sectors and the stocks listed in NIFTY 50 constitutes a future research work.

Acknowledgement

The authors express sincere gratitude to the faculty and staff of the Department of Artificial Intelligence and Data Science, Genba Sopanrao Moze College of Engineering Pune, for providing the necessary infrastructure and resources for this research. We acknowledge the valuable guidance received from our project mentors throughout the development of AI based Equity Portfolio Management System. Special appreciation is extended to the open-source community for making available the tools and libraries that facilitated this work. We also thank our colleagues for their constructive feedback during the development and testing phases of this platform.

References

1. Brim, A. (2020) “Deep reinforcement learning pairs trading with a double deep Q-network”,Proceedings of the 10th Annual Computing and Communication Workshop and Conference (CCWC’20), pp. 222-227
2. Carta, S. M., Consoll, S., Podda, A. S., Recupero, D. R. and Stanciu, M. M. (2021) “Ensembling and dynamic asset selection for risk-controlled statistical arbitrage”, IEEE Access, Vol. 9, pp. 29942-29959. DOI: 10.1109/ACCESS.2021.3059187.
3. Chatterjee, A., Bhowmick, H. and Sen, J. (2021) “Stock price prediction using time series, econometric, machine learning, and deep learning models”, Proceedings of the IEEE Mysore Sub Section International Conference (MysuruCon’21), pp. 289-296, October 24-25, Hassan, Karnataka, India. DOI: 10.1109/MysuruCon52639.2021.9641610.
4. Chen, Y-Y., Chen, W-L. and Huang, S-H. (2018) “Developing arbitrage strategy in high-frequency pairs trading with filterbank CNN algorithm”, Proceedings of the 2018 IEEE International Conference on Agents (ICA’18), pp. 113-116, July 28-31, 2018, Singapore. DOI: 10.1109/AGENTS.2018.8459920.
5. Cheng, D., Liu, Y., Niu, Z. and Zhang, L. (2018) “Modeling similarities among multi-dimensional financial time series”, IEEE Access, Vol. 6, pp. 43404 – 43414. DOI: 19.1109/ACCESS.2018.2862908.
6. Chong, E., Han, C. and Park, F. C. (2017) “Deep learning networks for stock market analysis and prediction: Methodology, data”, Experts Systems with Applications, Vol. 85, pp. 187-205, October 2017. DOI: 10.1016/j.eswa.2017.04.030.
7. Chou, Y-H., Jiang, Y-C. and Kuo, S-Y. (2021) “Portfolio optimization in both long and short selling trading using trend ratios and quantum-inspired evolutionary algorithms”, IEEE Access, Vol. 9, pp. 152115-152130. DOI: 10.1109/ACCESS.2021.3126652.
8. Corazza, M., di Tollo, G., Fasano, G. and Pesenti, R. (2021) “A novel hybrid PSO-based metaheuristic for costly portfolio selection problem”, Annals of Operations Research, Vol. 304, pp. 109-137. DOI: 10.1007/s10479-021-04075-3.
9. Fama, E. F. (1970) “Efficient capital markets: A review of theory and empirical work”, Journal of Finance, Vol. 25, No. 2, pp. 383-417. DOI: 10.2307/2325486.