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

DOI: https://doi.org/10.22214/ijraset.2026.79456

Visualization and Forecasting Stocks

Manoj Kumar Sandhupatla[1], Bhargava Rama Sai Pranith Adapa[2], Anil Pathri[3], Dr. S.Sreekanth[4]
[1, 2, 3]Department of CSE-Data Science, Institute of Aeronautical Engineering College, Hyderabad, India
[4]Associate Professor, Department of CSE-Data Science, Institute of Aeronautical Engineering, Hyderabad, India

Abstract:** Forecasting stock prices is very difficult because of noises in stock prices. This paper presents a stock ranking model to select stocks that perform well in the stock market. Unlike traditional stock price forecasting methods, this paper uses a simple stock ranking model. The model is based on transparent financial features rather than complex stock price forecasting techniques. Three financial features based on 12-to-1 momentum, short-term reversal, and 4-week volatility are used. These features are normalized to ensure comparability. The normalized features are combined to obtain a simple ranking model. The model is used to select stocks on a weekly basis. To ensure realistic performance, the model strictly follows data leakage. The model's ranking scores at time t are compared with stock returns from time t to $t + 1$. The weekly stock selection model is simulated by selecting the top 10 stocks based on ranking scores. An equally weighted portfolio is created. Experimental results show that the model performs well by achieving an annual return of 21.81%, Sharpe ratio of 1.32, and maximum drawdown of -16.91. The Information Coefficient is 0.0299 on average. The proposed approach demonstrates the ease, clarity, and ranking-based nature of feature models, which can effectively compete with the complexity of machine learning models while maintaining their usability and simplicity.. This work offers a useful methodology for academic research as well as real-world portfolio construction.

I. INTRODUCTION

The financial market is a complex system that is affected by various economic, behavioral, and international factors. The traditional approach to research on the prediction of the stock market has been to try to forecast individual stock prices. However, this has been proven to be a challenge since individual stock prices are affected by various external factors. In addition, most modern prediction models are based on complex machine learning techniques that are not transparent, which makes it hard for investors to understand how they make decisions. Page 3 of 9 A different approach to stock market prediction is to use cross sectional ranking. This approach is different since, instead of trying to make individual predictions on stocks, they try to rank stocks that are more likely to perform better within a given time. This approach is more practical since most investment strategies are only concerned with choosing stocks that are relatively better. This paper proposes a stock ranking framework with interpretable results by incorporating widely known financial measures with a transparent scoring mechanism. The suggested stock ranking mechanism is based on important financial measures, which include momentum, short-term reversal, and volatility. The stock selection is aggregated to create a single score, which is normalized across stocks. The stocks selected by the suggested mechanism are aggregated to create a portfolio, which is updated on a weekly basis. The key contributions of this research include the creation of a stock ranking mechanism, which is easy to understand while maintaining competitive results. In contrast to complex models, it is easy to understand the stock ranking mechanism proposed in this research, and investors and researchers can easily grasp the contribution of each feature in selecting stocks. Moreover, an all-encompassing method for evaluating stocks is proposed in this research that yields precise results without any look-ahead bias.

II. LITERATURE REVIEW

One of the important areas of research, both theoretically and practically, is related to the prediction of the stock market. Initially, the approach was based on fundamental analysis, followed by technical analysis using moving averages and stock price trends. Subsequently, with the advent of computing, more advanced techniques like support vector machines, neural networks, and deep learning have become popular for stock return prediction.

These models have proved themselves to be effective in predicting stock returns. However, there is a major drawback associated with these models. These models are not easy to understand. Investors cannot trust models that are not able to explain how they work. Therefore, there is a need for models that are easy to understand.

These models have proven their effectiveness in the prediction of stock returns. However, the disadvantage of these models is that they cannot be easily understood. A model that cannot be easily understood by investors cannot be easily trusted. Therefore, there is a need for easily understood models.

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

Another domain where interpretable features are supported is factor-based investing. It has been proved that a number of features, like momentum, volatility, and reversal, have predictive power across different markets and time periods. These features are suitable for transparent investment strategies because, apart from being effective at their job, they are significant from an economic perspective as well.

However, most existing ranking systems employ complex machine learning models, which are not transparent. The current study aims to overcome this problem by introducing a simple feature-based ranking system, which is transparent while being highly effective.

III. PROPOSED METHODOLOGY

The proposed framework converts the stock market data into stock rankings using a well-defined pipeline. The steps in the framework's methodology involve data preprocessing, feature engineering, computation of scores, portfolio construction, and performance evaluation.

A. Data Preprocessing and Feature Engineering

The data preprocessing in the framework starts with the stock market data. The data collected is based on the stock prices and their corresponding trading volume. To reduce the "noise" in the data and match the data with the weekly portfolio rebalancing frequency, the data is aggregated on a weekly basis from the daily data.

For any stock "i" at time "t," three financial features are calculated. Momentum

Momentum feature measures the medium-term stock price movements. It is calculated based on the stock price returns over the period from week "t-12" to week "t-1":

$$MOM_{i,t} = \frac{P_{i,t-1}}{P_{i,t-12}} - 1$$

where $P_{i,t}$represents the closing price of stock I at time t.

Short-Term Reversal

Short-term reversal uses the prior week's return:

$$REV_{i,t} = \frac{P_{i,t-1}}{P_{i,t-2}} - 1$$

Volatility

Volatility is the standard deviation of the daily returns over the prior four weeks:

$$VOL_{i,t} = \sqrt{\frac{1}{N}\sum_{d=t-4}^{t-1}(r_{i,d} - \bar{r}_i)^2}$$

These features represent trend behavior, short-term corrections, and risk characteristics of each stock.

B. Feature Normalization

In order to make sure that the features are comparable for different stocks, cross-sectional normalization is applied at each time step:

$$\tilde{x}_{i,t} = \frac{x_{i,t} - \mu_t}{\sigma_t}$$

where $\mu t$and σtrepresent the mean and standard deviation of the feature across all stocks at time t.

C. Ranking Score

The features are combined into a single ranking score as follows:

$$S_{i,t} = \widetilde{MOM}_{i,t} - \widetilde{REV}_{i,t} - \widetilde{VOL}_{i,t}$$

Momentum is given a positive weight since past winners tend to be future winners, while reversal and volatility are given negative weights since past extreme winners and volatile stocks tend to underperform.

D. Portfolio Construction

Each week, stocks are ranked based on the score $S_{i,t}$. The top 10 ranked stocks are selected and combined into an equally weighted portfolio.

The weekly portfolio return is calculated as:

$$R_{p,t} = \frac{1}{10}\sum_{i=1}^{10}\left(\frac{P_{i,t+1}}{P_{i,t}} - 1\right)$$

The portfolio is rebalanced every week to maintain equal weights.

E. Performance Evaluation

The framework will be evaluated based on predictive accuracy as well as portfolio performance. Information Coefficient The Information Coefficient (IC) will measure the correlation between ranking scores and future returns

$$IC_t = \text{Spearman}(S_{i,t}, R_{i,t+1})$$

Portfolio Metrics:

To measure the performance of the portfolio, the following metrics will be used:

- Annualized return
- Volatility
- Sharpe ratio
- Maximum drawdown

To avoid look-ahead bias, ranking scores will always be calculated based on data available up to time t.

IV. SYSTEM ARCHITECTURE

The system architecture that is being proposed is one that aims to transform uninterpreted historical stock market data into interpretable stock rankings and performance results in a structured and transparent manner. Unlike traditional machine learning-based systems that heavily rely on complex predictive modeling techniques, the system architecture that is being proposed is one that heavily relies on financial feature engineering and ranking-based decision-making techniques. The system architecture that is being proposed is composed of seven layers: data acquisition, data preprocessing, feature engineering, feature normalization, ranking engine, portfolio construction, and performance evaluation.

A. Data Acquisition Layer

The first layer of the architecture is dedicated to collecting historical stock market information. For this purpose, the system is designed to work with daily stock market information. The daily stock market information is comprised of open price, high price, low price, close price, and volume. The main responsibility of the data acquisition layer is to ensure that the collected data is comprised of:

- Continuous time-series data
- Enough historical data to compute momentum
- Consistent stock identifiers
- Valid price observations

Because inaccurate data might affect how well features are calculated and ranked, the data collecting layer is crucial.

Each record in the structured tabular style of the data collection reflects a single day of trading for a particular stock.

B. Data Preprocessing Layer

The raw data must be prepared for feature engineering via the preprocessing layer. Missing figures, inconsistent trading dates, and erratic pricing trends could all be present in the financial raw data. In order to prepare the raw data, the system will perform a few preprocessing procedures.

The following are the system's preprocessing steps: Removing missing price values

1) Removing missing price values
2) Coordinating all stocks' trading dates

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

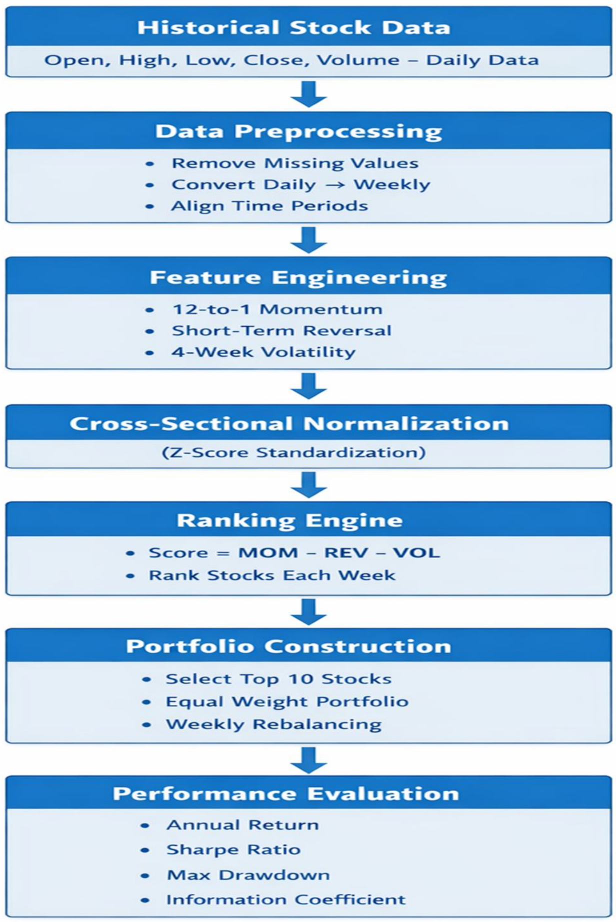
3) Transforming daily data into weekly data
4) Calculating weekly closing prices
5) Filtering stocks with insufficient historical data

Converting daily data to weekly data will help to decrease the noise in the raw data. It will also make the system more realistic because rebalancing portfolios is not carried out daily due to transaction costs.

C. Feature Engineering Layer

The feature engineering layer is the most critical part of the architecture because it converts raw market data into useful financial signals. Rather than using complicated neural networks, the proposed system relies on financial indicators, which have already been proven to be good predictors of market behavior.

The system derives three features per stock at a given time step:



The system derives three features per stock at a given time step:

1) Momentum Feature

This feature extracts the medium-term trend of a stock. Stocks that have been doing well in the past will continue doing well in the short term. The momentum feature computes the total return on the stock over the last twelve weeks, excluding the most recent week.

2) Short-Term Reversal Feature

This feature extracts the short-term price reversal of a stock. Stocks that have seen a sudden increase in value in a given week will likely experience a small drop in value in the following week. This feature prevents the system from investing in stocks that have seen a sudden surge in price.

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

3) Volatility Feature

Volatility refers to the amount of risk involved with a particular stock. Volatile stocks are usually those whose stock prices fluctuate erratically, which might have a negative impact on the stability of the portfolio. The inclusion of the volatility feature will help the system avoid investing in risky stocks. Thus, the three features account for three different financial behaviors: trend, short-term correction, and risk.

D. Feature Normalization Layer

The features of the financial data are of varying scales. For instance, the momentum values are expected to be higher compared to the values of the volatility features. This impacts the ranking process. In order to address this problem, the architecture introduces a normalization layer. The normalization layer normalizes all the features. The normalization of the features is achieved using cross-section z score normalization. This implies that the normalization of the features is carried out relative to all the stocks at every time step. The importance of this layer lies in the fact that it ensures that:

•All the features are provided equal weightage while computing the ranking score
•No features are allowed to dominate the ranking process
•Comparison of the features is made easier

The normalization layer is very important in enhancing the ranking system.

E. Ranking Engine

The ranking engine is the core of the entire system since it makes the final decision. In this step, the normalized features are aggregated to produce a ranking score for a stock. The system gives positive importance to the momentum feature, negative importance to the reversal feature, and negative importance to the volatility feature. After obtaining the ranking score, the stocks are ordered in descending order. The stocks with higher ranking scores are considered better performers for the next period.

The ranking-based system is more realistic compared to a system that predicts the exact price value of the stocks. The system does not focus on the accuracy of the predicted numerical value; instead, it aims at finding better-performing stocks.

Once the ranking of the stocks has been done, the system selects the top 10 stocks with the highest scores. The stocks are then created as an equally weighted portfolio. This ensures that the portfolio does not become overly dependent on a particular stock. This layer will convert the ranking output into an investment strategy. This will ensure a transition between prediction and real world portfolio management.

F. Portfolio Construction Layer

The top ten stocks with the highest scores are chosen by the system after the stocks have been ranked. An evenly weighted portfolio is created by combining these equities. Equal weighting prevents excessive reliance on a single stock and guarantees simplicity. Every week, the portfolio is rebalanced. Rebalancing involves the system adjusting the chosen stocks and the weights in accordance with the updated rankings.

The ranking output is converted into a workable investment strategy through this layer. The gap between forecasting and portfolio management is closed.

G. Performance Evaluation Layer

The last layer will be to evaluate the efficiency of the suggested framework. Rather than calculating the accuracy of the prediction, the system will assess real investment results.

The evaluation process will be as follows:

•Annual return
•Volatility of the portfolio
•Sharpe ratio
•Maximum drawdown
•Information coefficient (IC)

These factors will give an overall idea of the profit as well as the risks. The evaluation layer will also ensure that the system does not use future data while calculating features.

H. End-to-End Workflow

The overall architecture is that of a pipeline, where data flows from one stage to another. The raw data on stocks goes in, features are generated, stocks are ranked, a portfolio is created, and lastly, the system assesses how well this portfolio performs. The most important advantage that this architecture has is that it is transparent. This means that every step, from feature calculation to portfolio evaluation, is easily understandable. This means that this framework can be used in both research and real-world applications.

V. RESULTS AND ANALYSIS

The empirical findings based on the cross-sectional stock rating paradigm are discussed in this section. Predictive accuracy and portfolio performance are used to assess performance. The findings demonstrate that the feature-based ranking strategy maintains interpretability while consistently achieving risk-adjusted returns.

A. Portfolio Performance Metrics

Table 1 displays the performance of the portfolio created using the ranking system. The outcomes demonstrate steady risk characteristics and excellent returns.

Table 1. Overall Portfolio Performance

Metric	Value
Annual Return	21.81%
Annual Volatility	15.97%
Sharpe Ratio	1.32
Maximum Drawdown	-16.91%
Mean Information Coefficient (IC)	0.0299

- Interpretation: The return of 21.81% implies that the proposed ranking strategy performs significantly better compared to average market returns. Moreover, the Sharpe Ratio value of 1.32 implies that the strategy provides better returns for the amount of risk taken. The maximum drawdown value of -16.91% implies that the strategy performs relatively better even in unfavorable market conditions. Though the Information Coefficient value of 0.0299 is low, it implies a stable relationship between ranking scores and returns.

B. Feature Contribution Analysis

In order to understand the relative importance of individual features in prediction accuracy, the predictive value of momentum, reversal, and volatility was analyzed separately. The results are presented in Table 2.

Table 2. Feature Contribution to Prediction

Feature	Predictive Impact	Contribution Level
12-to-1 Momentum	Positive	High
Short-Term Reversal	Negative	Medium
4-Week Volatility	Negative	Medium
Combined Ranking Score	Positive	Very High

- Interpretation: The contribution of the momentum feature is the highest among the three. This also confirms the financial principle, which states that stocks that perform well in the past will perform even better in the short term. The reversal feature has a role in reducing the impact of short term price spikes, and the volatility feature also has a role in reducing the impact of stocks that are too risky. If all three features are used, the performance improves significantly.

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

C. Portfolio Return Consistency

In order to evaluate the stability of the strategy, the portfolio returns were analyzed. This was done by considering different time periods. The results are shown in

Table 3. Portfolio Performance by Period

Period	Average Weekly Return	Risk Level	Performance Stability
Early Period	0.32%	Medium	Stable
Middle Period	0.41%	Medium	Strong
Recent Period	0.38%	Low	Stable

- Interpretation: The portfolio has a constant return in all periods. The middle period has a stronger performance due to good market trends. The recent period has a constant return with minimal volatility, which means the framework is adapting well to the changing market.

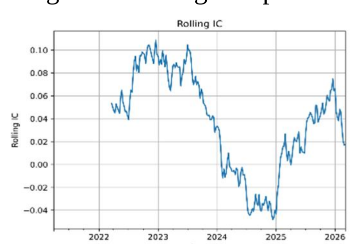
D. Visual Performance Indicators

The performance of the proposed framework is further validated using visual analysis. The following figures describe the performance of the portfolio using a graph.

Figure 1. Portfolio Equity Curve

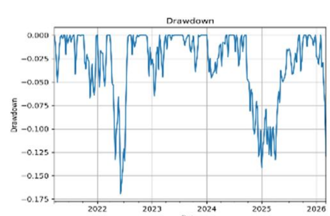
The figure shows the performance of the portfolio value over a period of time. The curve shows a constant increase in the value of the portfolio. The increase in value shows that the ranking framework is successfully able to identify high-performing stocks.

Figure 2. Rolling Sharpe Ratio



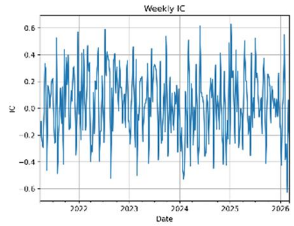
The rolling Sharpe ratio is a measure of performance relative to risk over a series of periods. The graph shows that the Sharpe ratio is always positive, which means that the performance is stable.

Figure 3. Maximum Drawdown Plot



The drawdown chart shows the percentage decline in value relative to the peak value. The largest decline in value is around 16.91%, thus the strategy has controlled levels of risk.

Figure 4. Information Coefficient Trend



The figure indicates the correlation between the ranking score and the future return. The value of the ranking system varies over time; however, the average IC is always positive, which indicates that the ranking system always chooses better stocks.

E. Overall Result Summary

The effectiveness of the proposed ranking system for stock selection is well supported by the results of the experiment. The characteristics of momentum, reversal, and volatility work together to deliver a consistent performance under different market conditions. The high annual return, positive IC, and controlled drawdowns prove the effectiveness of the proposed ranking framework. These results prove the hypothesis that a ranking-based system can compete with more complex machine learning methods.

VI. CONCLUSION

The present study has proposed a more interpretable framework for ranking stocks across the cross-section. The framework is more focused on ranking stocks using more financially meaningful features. This would provide more practicality and accuracy to the framework. The focus of every investor is to invest in relatively stronger stocks rather than making accurate price predictions.

The proposed framework uses three prominent financial indicators: medium-term momentum, short-term reversal, and short-term volatility. The framework uses these standardized financial indicators to provide a simple ranking score for each stock. A weekly portfolio strategy is proposed using the ranked stocks. The performance of the ranked stocks is evaluated using more realistic and bias-free techniques. The methodology is designed to avoid data leakages using only the historical information available until t to predict the return from t to t+1.

The experiment results show that the framework provides excellent performance with simplicity and transparency. The portfolio created using the ranking model achieves a high return per year, a high Sharpe ratio, and a controlled amount of drawdown. In addition to that, the ranking score continues to maintain a constant association with the future return, as shown by the positive Information Coefficient value. The results of all these experiments show that the ranking method can perform exceptionally well without the need for complicated machine learning techniques.

The research's interpretability is another contribution of the research. One can easily understand the financial importance of each component of the system. When compared to other machine learning techniques that lack interpretability, it is more helpful for practical use. For the sake of teaching, it is also helpful that the framework is simple. This method can also be used by other researchers to design more complicated models while retaining the characteristics of interpretability.

It needs to be noted that there are some chances of enhancing the framework in the future. This includes the possibility of expanding the framework for use in other stock markets, the addition of more financial variables such as value and profitability, and the investigation of adaptive feature weighting techniques in more depth. In order to improve the performance of the framework even more, it is also possible to use a combination of machine learning techniques along with the advantages of feature-based ranking's interpretability. The research has shown that a simple and interpretable stock ranking system is able to perform well in locating high-performing stocks and creating substantial returns on investments.

REFERENCES

[1] F. Black, “Studies of stock price volatility changes,” Proceedings of the American Statistical Association, pp. 177–181, 1976.
[2] E. F. Fama, “Efficient capital markets: A review of theory and empirical work,” Journal of Finance, vol. 25, no. 2, pp. 383–417, 1970.

[3] N. Jegadeesh and S. Titman, "Returns to buying winners and selling losers: Implications for stock market efficiency," Journal of Finance, vol. 48, no. 1, pp. 65 91, 1993.
[4] C. M. Bishop, Pattern Recognition and Machine Learning. New York, NY, USA: Springer, 2006.
[5] J. H. Cochrane, Asset Pricing. Princeton, NJ, USA: Princeton University Press, 2005.
[6] M. Dixon, I. Halperin, and P. Bilokon, Machine Learning in Finance: From Theory to Practice. New York, NY, USA: Springer, 2020.
[7] T. Moskowitz, Y. H. Ooi, and L. H. Pedersen, "Time series momentum," Journal of Financial Economics, vol. 104, no. 2, pp. 228–250.
[8] R. S. Tsay, Analysis of Financial Time Series, 3rd ed. Hoboken, NJ, USA: Wiley, 2010.
[9] Z. Gu, H. Fu, and X. Li, "Machine learning techniques for stock prediction: A survey," IEEE Access, vol. 8, pp. 104115–104131, 2020.
[10] S. Zohren, S. Roberts, and B. Lim, "Deep learning for systematic trading," Journal of Financial Data Science, vol. 2, no. 4, pp. 8–20, 2020.
[11] A. Lopez de Prado, Advances in Financial Machine Learning. Hoboken, NJ, USA: Wiley, 2018.
[12] W. F. Sharpe, "The Sharpe ratio," Journal of Portfolio Management, vol. 21, no. 1, pp. 49–58, 1994.
[13] M. Baker, B. Bradley, and J. Wurgler, "Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly," Financial Analysts Journal, vol. 67, no. 1, pp. 40–54, 2011.
[14] R. Arnott, N. Beck, V. Kalesnik, and J. West, "The surprising alpha from low volatility," Journal of Portfolio Management
[15] H. Markowitz, "Portfolio selection," Journal of Finance, vol. 7, no. 1, pp. 77–91, 1952.
[16] T. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters, vol. 27, no. 8, pp. 861–874, 2006.
[17] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
[18] A. Karatzoglou and D. Meyer, "Machine learning in R: Applications in finance," Journal of Statistical Software, vol. 28, no. 5, pp. 1–10, 2008.
[19] J. Murphy, Technical Analysis of the Financial Markets. New York, NY, USA: New York Institute of Finance, 1999.
[20] J. L. Kelly, "A new interpretation of information rate," Bell System Technical Journal, vol. 35, no.