

Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting

Santosh S. Shinde ^{1,✉}, Jagannath E. Nalavade ¹

1. Department of Computer Science and Engineering, The MIT School of Computing, MIT Art, Design and Technology University, Pune, IND

Received: June 19, 2026 | Review began: July 16, 2026 | Review ended: October 22, 2026 | Published: October 25, 2026

© Copyright 2026

This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Stock market analytics is a vital part of systematic investment decision-making, portfolio management, and risk management. Conventional price-based evaluation procedures do not consider volatility, downside risk, and liquidity limitations, resulting in inappropriate stock selection. This paper suggests a combined risk-adjusted analytical framework that is applied to the constituents of the NIFTY 50 index based on historical data from 2015 to 2025. The dataset contains historical data for 52 unique companies that were constituents of the NIFTY 50 index during the study period (2015-2025). Although the index consists of 50 companies at any point in time, constituent changes over the study period resulted in a total of 52 unique companies being included in the analysis. This approach is a combination of the Compounded Annual Growth Rate (CAGR), annualized volatility, Sharpe ratio, Sortino ratio, and Maximum Drawdown to create a composite performance score. To ensure feasible tradability, liquidity screening is included. K-means clustering ($K = 3$) is used to carry out structural market segmentation based on growth and risk dimensions, and predictive power is determined through AutoRegressive Integrated Moving Average-based time-series forecasting. According to the empirical evidence, the top five risk-adjusted performers include TITAN, HINDUNILVR, BAJFINANCE, KOTAKBANK, and HDFC, with composite scores ranging from 0.833 to 1.829. The clustering results show three different performance regimes, with high-CAGR stocks (~26%) exhibiting better Sharpe ratios (~0.66) and lower volatility (~32%), while underperforming groups show negative growth and extreme drawdowns (-89%). The research adds a unified and quantitatively reliable stock assessment paradigm that balances growth, risk, and liquidity. In practice, the method aids robust stock ranking and provides a foundation for improving portfolio allocation strategies and predictive modeling.

Categories: AI/ML-based decision support systems, Algorithm Analysis, Computational Science and Engineering

Keywords: risk-adjusted return, sharpe ratio, k-means clustering, arima forecasting, nifty 50

Introduction

The consequence of financial markets is the generation of high-frequency, high-volume, and heterogeneous data, where actionable information is now becoming systematic and data-driven, as opposed to decision-making that is more intuitive. With indexes such as NIFTY 50, the challenge is not only to understand how prices move but also how to turn this multi-dimensional market data into reliable ranking, grouping, and prediction pipelines that can still be clearly understood by investors and utilized in portfolio construction. Recent work has therefore moved toward learning-based and hybrid approaches to decision-making, which combine financial theory with high-quality machine learning for both portfolio optimization and trading policy learning under uncertainty and non-stationarity [1].

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

To build momentum portfolios that have a higher implementation capability, Aslam [1] proposed clustering-based stock selection on the basis of multiple risk-adjusted aspects. To capture greater resilience to uncertainty in the market, Yang [2] presented a deep reinforcement learning (DRL) framework for portfolio management that learns dynamic allocation strategies from market observations and dynamically adjusts portfolio weights under changing market conditions. Choi and Kim [3] proposed an expert-informed DRL approach for portfolio optimization that incorporates domain knowledge into the learning process. Cui et al. [4] introduced a DRL hyper-heuristic framework for multi-period portfolio optimization to address action-space complexity and improve sequential investment decisions. Jang and Seong [5] proposed a DRL model for stock portfolio optimization that balances expected return and risk. Zhang [6] proposed a DRL model for stock portfolio optimization that balances risk and return within the framework of modern portfolio theory.

Sun et al. [7] used GraphSAGE with DRL to learn non-Euclidean graphs between indices and stocks with the aim of optimizing a portfolio. The article by Grudniewicz and Ślepaczuk [8] touched on the area of machine learning-based asset allocation, wherein emphasis is placed on robustness and out-of-sample performance in an unstable market. Soleymani and Paquet [9] proposed DeepPocket (graph convolution + DRL) to enhance feature learning for financial portfolio management. Soleymani and Paquet [10] proposed DeepBreath (limited stacked autoencoder + DRL) for online portfolio optimization. Jiang et al. [11] explored DRL in portfolio choice with continuous actions in a multi-dimensional state space. In a move to improve the stability of trading, de Azevedo Takara et al. [12] proposed extending trading deep Q-network (DQN) to address the sparsity of intraday trading rewards. Shavandi and Khedmati [13] proposed employing a multi-agent DRL trading structure in which multiple time-horizon agents are employed, and risk is used to examine performance.

A multi-factor stock trading model developed by Liu et al. [14] relied on a combination of DQN and sequence feature extractors/attention systems. A graph-based concept of learning was applied by Ekmekcioğlu and Pınar [15], in which the relational structure was incorporated into the dynamics of assets. In order to develop decision-making policies through actions based on price sequences, Shi et al. [16] proposed stock trading rule discovery through the application of DQNs twice. Cao et al. [17] looked at more advanced trading algorithms based on machine learning, with special attention to the problem of robustness and evaluation design in financial prediction. Jung and Oh [18] provided a feature engineering and stability-based study of stock market predictors using quantitative finance. Wu et al. [19] performed a survey of machine learning-based stock prediction methods and emphasized the importance of multi-metric evaluation. Separately, Sezer et al. [20] presented a systematic review of deep learning techniques for financial time-series forecasting.

Price-only research, that is, inferential research based on close/open price action analysis or direction forecasting but not on sufficient consideration of volatility, downside risk, liquidity, and sensitivity to drawdowns, has been a longstanding weakness in the majority of applied stock research. Price signals can be deceptively optimistic: they may concentrate on high raw-return assets, in addition to being misleading indicators of inadequate risk coverage, inconsistent return distributions, or poor return tradability. This variance is amplified in real index constituents, where two stocks with comparable cumulative returns do not have comparable downside profiles or liquidity positions, and this has a direct bearing on implementability, transaction costs, and achieved performance. In this regard, risk-sensitive pipelines employ more risk-adjusted analysis and selection instead of return-based screening, which is applied in a risk-blind environment [1].

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. *Cureus J Comput Sci* 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>



FIGURE 1: Research Framework for NIFTY 50 Analytics

To address these practical shortcomings, recent works emphasize risk-adjusted ranking and decision-making, often with a combination of multiple performance measures (e.g., Sharpe/Sortino-type measures and drawdown-sensitive scoring), clustering, and allocation policies [1]. Figure 7 presents the proposed research framework for NIFTY 50 analytics, including data extraction, cleaning, financial metric computation, liquidity filtering, composite scoring, and stock clustering. As an example, evidence from the literature on portfolio optimization and policy learning shows that maximizing raw returns per se is insufficient; effective systems include the goal of maximizing returns and minimizing volatility, downside exposure, and changes of regime [2]. This promotes the active work to consider the issue of stock selection and grouping as a sequential financial analytics problem, i.e., beginning with clean data and valid metrics, followed by the construction of the composite scores to rank and fill with clustering to respond to the structure among the constitutes, and finally, mimicking forecasting to support decision schedules. The paper proposes a minimal, concise analytical framework of NIFTY 50 analytics that (i) converts cleaned Open, High, Low, Close, and Volume data into financial measures that are simple to comprehend, (ii) introduces a liquidity-sensitive filter to discourage unrealistic decisions, (iii) forms a composite (risk-adjusted) score to rank the stocks beyond the price-only reasoning, (iv) clusters (defines groups with similar risk-return behaviors) to draw comparative and diversification ideas, and (v) incorporates forecasting models for representative stocks from selected clusters and top-ranked stocks to support near-term market trend analysis and informed trading decisions.

Materials And Methods

The methodology used for the risk-adjusted performance evaluation, liquidity screening, clustering analysis, and time-series forecasting of NIFTY 50 constituent stocks is described. The proposed framework spans six steps: (i) data processing, (ii) calculation of financial performance indicators, (iii) liquidity screening, (iv) composite risk-adjusted scoring and ranking, (v) K-means clustering, and (vi) time-series forecasting. The methodology is designed to guarantee statistical validity, financial interpretability, and applicability in investment decision-making.

Data preprocessing

The historical stock data were downloaded from individual company-level CSV files that provide daily trading data for NIFTY 50 constituent companies. All the company-specific files were combined into a single file to enable cross-sectional analysis and evaluate performance at the portfolio level. To ensure structural consistency, index-level files (NIFTY50_all.csv) and metadata files were excluded. All records were converted to the standard datetime format and then sorted by symbol and date for the proper calculation of returns and rolling financial metrics. These analyses were limited to the last 10 years to ensure consistency in the analysis and comparison of all companies. Missing observations were

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

eliminated instead of being imputed, as risk-adjusted performance measures such as volatility, the Sharpe ratio, and the Sortino ratio are sensitive to incomplete time-series data. In addition, whitespace was removed from the stock symbols, and only the equity (EQ) series were retained to ensure data homogeneity. The financial performance indicators were calculated after preprocessing, and the return, risk, downside risk, and long-term growth characteristics of each stock were evaluated. The daily return is the percentage change in the closing price over two consecutive trading days and can be computed as shown in Equation (1):

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (1)$$

where

R_t = Daily return at time t

P_t = Closing price on day t

P_{t-1} = Closing price on the previous trading day

The Compound Annual Growth Rate (CAGR) is the average annual growth rate of an investment over the study period and is calculated using Equation (2). It denotes the closing price at time t .

$$\text{CAGR} = \left(\frac{V_f}{V_i} \right)^{\frac{1}{n}} - 1 \quad (2)$$

where

V_f = Final investment value

V_i = Initial investment value

n = Investment period (years)

Annualized volatility measures the variability of stock returns and is calculated from the standard deviation of daily returns, as shown in Equation (3).

$$\sigma_{\text{annual}} = \sigma_{\text{daily}} \sqrt{T} \quad (3)$$

where

σ_{annual} = Annualized volatility

σ_{daily} = Standard deviation of daily returns

T = Number of trading days per year (typically 252)

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (R_i - \bar{R})^2} \quad (4)$$

where

R_i = Daily return of the i^{th} observation

$\bar{R}$ = Average daily return

N = Total number of observations

The Sharpe ratio measures the excess return earned per unit of total risk and is calculated using Equation (5).

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

$$\text{Sharpe} = \frac{\text{CAGR} - R_f}{\sigma} \quad (5)$$

where

R_f = Annual risk-free interest rate

σ = Annualized standard deviation of stock returns (volatility)

The Sortino ratio measures the excess return relative to downside risk and is calculated using Equation (6).

$$\text{Sortino} = \frac{\text{CAGR} - R_f}{\sigma_d} \quad (6)$$

where

R_f = Risk-free rate

σ_d = Downside deviation

CAGR = Compound Annual Growth Rate

Maximum Drawdown (MDD) measures the largest percentage decline from a historical peak to a subsequent trough during the investment period and is calculated using Equation (7).

$$\text{MDD} = \min \left(\frac{V_t - \text{Peak}_t}{\text{Peak}_t} \right) \quad (7)$$

where

V_t = Portfolio value at time t

Peak_t = Highest portfolio value achieved up to time t

MDD measures the largest percentage decline from a historical peak to a subsequent trough during the investment period. Combined, these financial indicators measure stock performance in terms of return, growth, volatility, downside risk, and maximum loss of capital, providing a more complete picture by which to evaluate stock performance.

To ensure practical tradability, stocks were subjected to a two-stage liquidity screening process based on average daily trading volume and turnover. Securities below the 20th percentile in either metric were excluded to eliminate structurally illiquid stocks. The stocks had to meet the minimum requirements for average daily volume and turnover. This provides minimum liquidity and does not allow the inclusion of thinly traded securities. Stocks with a liquidity distribution below the 20th percentile for either volume or turnover were eliminated. This action removes structurally illiquid stocks, although they may meet absolute thresholds. The overall filtering mechanism ensures that the selected stocks are both operationally viable and statistically appealing. Because the performance metrics are based on different scales, they need to be normalized before they are aggregated. All the metrics were converted into Z-scores, which were calculated as shown in Equation (8):

$$Z = \frac{X - \mu}{\sigma} \quad (8)$$

where

X = Metric value

μ = Mean of the metric

σ = Standard deviation of the metric

The composite risk-adjusted score was calculated using Equation (9).

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

$$\text{Score} = 0.45Z_{\text{Sharpe}} + 0.35Z_{\text{CAGR}} + 0.15Z_{\text{Sortino}} + 0.05Z_{\text{Drawdown}} \quad (9)$$

where

Z_{Sharpe} = Normalized Sharpe ratio

Z_{CAGR} = Normalized Compound Annual Growth Rate

Z_{Sortino} = Normalized Sortino ratio

Z_{Drawdown} = Normalized Maximum Drawdown

The weighting scheme assigns the highest importance to the Sharpe ratio (45%), followed by CAGR (35%), the Sortino ratio (15%), and MDD (5%). The weighting gives greater emphasis to risk-adjusted performance (Sharpe) without diminishing the focus on growth (CAGR). The Sortino ratio provides additional downside sensitivity, while drawdown serves as a penalty for historically extreme losses. The weights were chosen so that they do not overemphasize any one dimension. Stocks were ranked in descending order of the composite score to determine the best-performing candidates.

Clustering method

K-means clustering was used to reveal the structural relationships between stocks using a few financial metrics. The number of optimal clusters was identified through the elbow method, according to which the within-cluster sum of squares (WCSS) is plotted against different values of K. The elbow point was observed at K = 3, indicating the best trade-off between clustering performance and model simplicity. The resulting clusters represent high-return/low-risk, moderate-return/moderate-risk, and low-return/high-risk stocks, thereby supporting portfolio diversification and investment decision-making. This segmentation facilitates structural classification rather than individual ranking, and the segmentation will help in developing portfolio diversification strategies. Figure 2 illustrates the elbow plot for determining the optimal number of clusters (K) using the WCSS. The elbow point observed at K = 3 indicates the optimal number of clusters for the K-means clustering analysis.

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

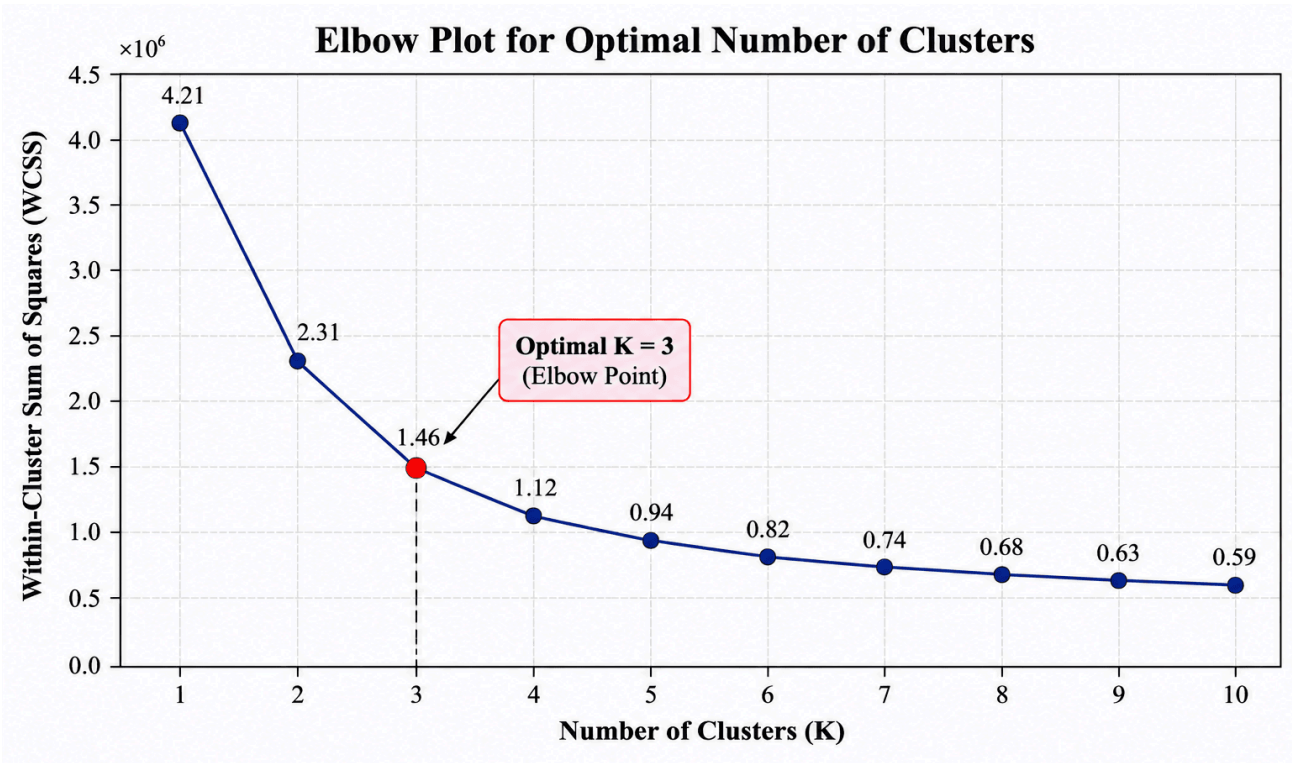


FIGURE 2: Elbow Method for Selecting the Optimal Number of Clusters

Results And Discussion

Experimental setup

The proposed framework was implemented in the following development environment: Python 3.11 and Visual Studio Code. The libraries used for data preprocessing and statistical analysis were Pandas and NumPy. The numerical analysis was performed using Python, and various financial performance indicators were calculated, such as daily return, CAGR, annualized volatility, the Sharpe ratio, the Sortino ratio, and MDD. Matplotlib and Seaborn were used for data visualization, and the optimal number of clusters was found using the elbow method with Scikit-learn and K-means clustering. The historical daily stock price data of NIFTY 50 constituent companies covering the period from 2015 to 2025 were used in the experimental analysis. The dataset consisted of Open, High, Low, Close, Previous Close, Last Price, Volume, and Turnover for each day. The dataset was preprocessed to remove missing data, standardize the stock symbols, convert the date data to datetime, and retain only EQ data prior to analysis. The experiments were performed on a computer with an Intel Core i5 processor, 8 GB of RAM, and the Windows 11 operating system. The proposed methodology was assessed using composite risk-adjusted scoring, liquidity screening, K-means clustering, and time-series forecasting using the AutoRegressive Integrated Moving Average (ARIMA) model. ARIMA is a statistical forecasting approach designed for univariate time-series data and is represented as $ARIMA(p, d, q)$, where p denotes the autoregressive order, d represents the degree of differencing required to achieve stationarity, and q denotes the moving-average order. In this study, ARIMA was employed to model the temporal behavior of the selected stock price series and generate forecasts of future prices based on their historical observations.

This section presents the research results based on the hypothesized analytical framework in stages. These findings are also discussed from statistical and monetary perspectives, with an emphasis on strength, structural compatibility, and investment relevance. The analysis is conducted in four phases, namely, (i) risk-adjusted ranking results, (ii) liquidity checking, (iii) structural segmentation through clustering, and (iv) evaluation of price stability.

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

Risk-adjusted ranking results

The composite scoring model combines the Sharpe ratio, CAGR, the Sortino ratio, and MDD by assigning a standardized weight to each. Ranking was done following liquidity screening, followed by Z-score normalization to maintain scale invariance across metrics. Table 1 highlights that TITAN and HINDUNILVR achieved high composite scores due to a balanced combination of strong CAGR and favorable risk-adjusted metrics, whereas BAJFINANCE, despite its high returns, exhibited higher volatility and drawdown.

Stock	CAGR	Volatility	Sharpe	Sortino	MDD	Composite Score
TITAN	0.255	0.337	0.564	0.818	-0.417	1.829
HINDUNILVR	0.207	0.240	0.590	1.021	-0.251	1.735
BAJFINANCE	0.264	0.479	0.416	0.398	-0.933	1.591
KOTAKBANK	0.157	0.327	0.281	0.303	-0.584	0.853
HDFC	0.149	0.289	0.289	0.416	-0.398	0.833

TABLE 1: Top 5 Stocks Based on Composite Score
CAGR, Compound Annual Growth Rate; MDD, Maximum Drawdown

The ranking shows a significant difference in growth-risk trade-offs. TITAN has the best composite score because of balanced results across all dimensions. Its Sharpe ratio (0.564) and Sortino ratio (0.818) show that it has efficiently compensated for total risk and downside risk. The MDD (-41.7%) indicates resilience during market downturns. HINDUNILVR has lower volatility (0.240) and the highest Sortino ratio (1.021), which indicates greater downside protection. Its CAGR is slightly lower than that of TITAN and BAJFINANCE, but its downside stability is higher, which increases its composite ranking. BAJFINANCE has the highest CAGR (0.264), as well as the highest volatility (0.479) and the largest drawdown (-93.3%). This high risk exposure is penalized by the composite score, demonstrating the application of risk-adjusted weighting. KOTAKBANK and HDFC demonstrate moderate growth with controlled volatility, which ranks them at the second level of performance. These findings validate the view that growth is not the determining factor in ranking position; rather, the composite model incentivizes long-term risk-adjusted efficiency. The stability of the ranking, from a methodological viewpoint, supports the soundness of the normalization and weighting process.

Liquidity validation

The liquidity measures were examined to ensure that the best-rated stocks can be traded in practice. Average daily volume and turnover values are presented in Table 2. Each of the five stocks has surpassed the absolute and relative liquidity levels. The turnover of HDFC and BAJFINANCE has been especially high, which indicates high institutional participation and capital inflows. The liquidity check ensures that the ranking model does not select illiquid but statistically attractive stocks. This step provides implementability and reduces execution risk in the application of portfolio allocation in the real world. As shown in Table 2, HDFC and BAJFINANCE exhibit the highest average turnover, indicating strong market participation and liquidity among the selected stocks.

Stock	Avg Volume	Avg Turnover (₹)
TITAN	2,076,269	1,426,772,766
HINDUNILVR	1,831,537	2,290,470,398
BAJFINANCE	1,360,945	4,050,095,197
KOTAKBANK	2,113,138	2,572,481,014
HDFC	3,301,651	4,848,475,984

TABLE 2: Liquidity Metrics of Top 5 Stocks

Clustering insights

The correlation between CAGR and other performance indicators was analyzed with the help of K-means clustering (K = 3). These findings are graphically summarized in Figure 3, which shows four bivariate projections: CAGR against annualized volatility, Sharpe ratio, maximum drawdown, and cumulative growth. The centers of the clusters are marked by red stars, and the small points represent the observation points of the stocks. The results of the clustering show a distinct division in terms of growth, risk, and wealth-creation levels.

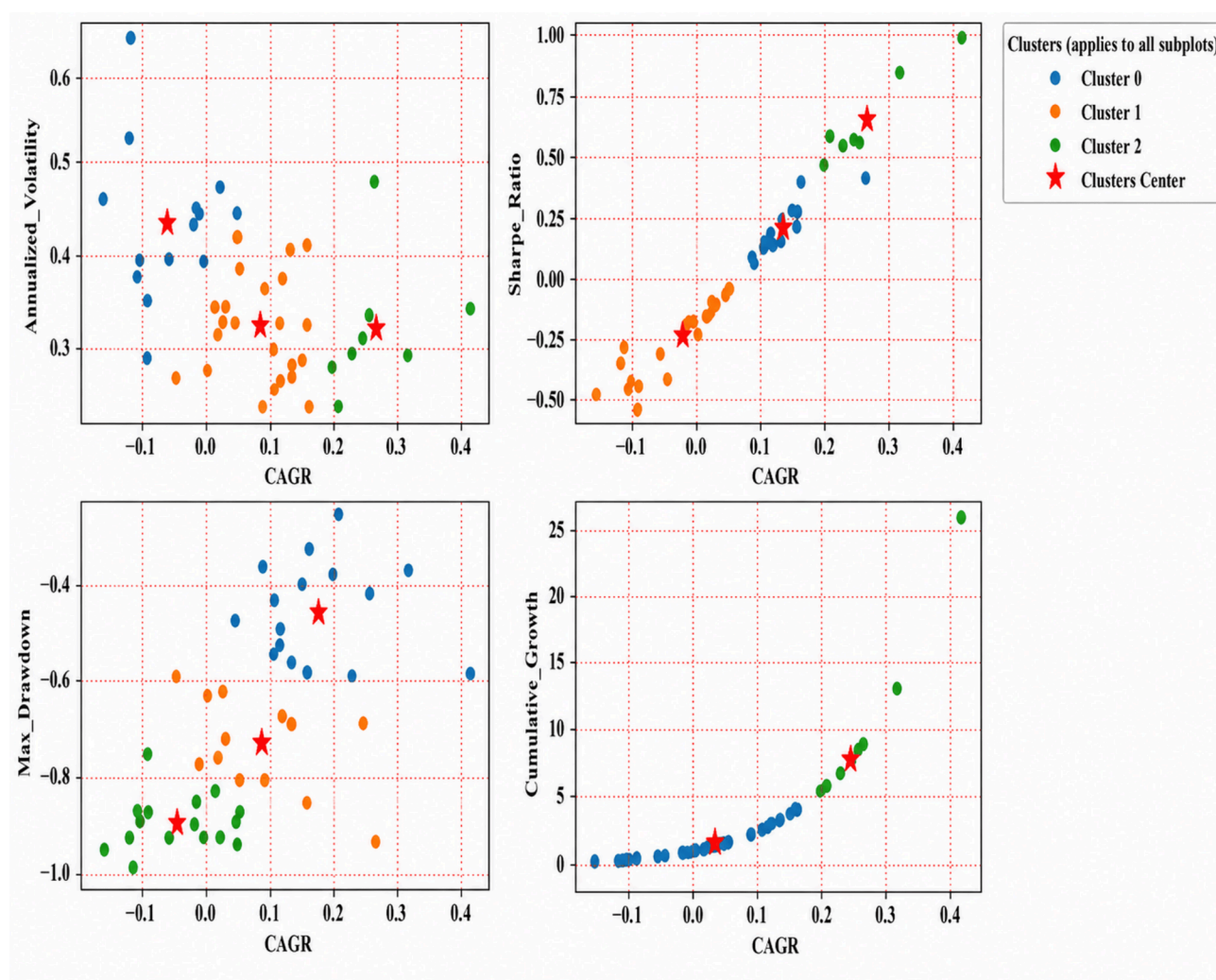


FIGURE 3: CAGR Versus Other Performance Parameters

CAGR, Compound Annual Growth Rate

As shown in Figure 3 (upper-left panel in the CAGR-volatility projection), there are three regimes. The first cluster, with the center (-0.0588, 0.4341), consists of stocks with slightly negative long-term growth and the highest annualized volatility (43.41%). These securities have fluctuating prices without compensating returns. The second cluster, with the center (0.0829, 0.3255), has moderate growth (8.29%) and managed volatility (32.55%), representing the transitional or mid-range performers. The third cluster is located at (0.2658, 0.3232) and has the highest CAGR (26.58%) with similar volatility to the mid-tier cluster (32.32%), which means that better growth is attained without being excessively volatile. The centroid alignment indicates that there is no positive correlation between long-term growth in the NIFTY 50 universe and volatility; instead, there is controlled dispersion in the highest-growth cluster.

As shown in the CAGR-Sharpe ratio projection (upper-right panel), there is a strong positive structural relationship between geometric growth and performance (risk-adjusted performance). The poorest-performing cluster is located at (-0.0246, -0.2283), indicating negative growth (-2.46%) and a negative Sharpe ratio (-0.2283), which represents an inefficient use of risk. The mid-level cluster is located at (0.1347, 0.2131), indicating moderate CAGR (13.47%) and a moderate positive Sharpe ratio (0.2131), suggesting limited yet acceptable risk compensation. The most effective cluster

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

has a center at (0.2660, 0.6595), with a high CAGR (26.60%) and a significantly better Sharpe ratio (0.6595). This correlates with the fact that increased long-term growth implies better risk-adjusted efficiency. The emphasis on the Sharpe ratio in the composite scoring system is therefore supported by the clustering structure.

Downside sensitivity is further demonstrated by the relationship between CAGR and MDD (lower-left panel). The cluster with the lowest values, (-0.0438, -0.8932), has negative growth (-4.38%) and a high drawdown (-89.32%), which indicates massive losses from the peaks and high fragility. The second group, located at (0.0860, -0.7319), is indicative of moderate positive growth (8.60%) but large drawdowns (-73.19%), implying uncertain wealth accumulation. The largest cluster, with a point at (0.1743, -0.4562), has a relatively higher CAGR (17.43%) and controlled drawdown (-45.62%). The fact that CAGR decreases monotonically with the extent of the drawdown suggests that extreme historical losses are structurally linked with lower long-term growth. This empirical trend justifies the inclusion of MDD as a penalty factor in the composite score.

The last forecast of CAGR/cumulative growth (lower-right panel) demonstrates the compounding impact of sustained returns per year. The low-growth cluster is oriented around (0.0307, 1.76), which entails a small rate of growth (3.07%) and a small cumulative growth in wealth (1.76 $\times$). The median group is (0.2446, 8.08), representing moderate-to-high CAGR (24.46%) that would translate into 8.08 $\times$ cumulative growth. The best-performing cluster, with a center of (0.4143, 26.29), has an exceptionally high CAGR (41.43%) and cumulative growth exponent (26.29 $\times$). The distortion of these clusters evidences the nonlinear amplification, or compounding, effect, whereby rising increments in CAGR cause a disproportionate increase in wealth accumulation over long periods of time.

In all the projections in Figure 3, three common structural regimes can be seen: a high-efficiency growth cluster with strong CAGR, a high Sharpe ratio, moderate volatility, controlled drawdown, and cumulative growth; a transient cluster with moderate performance on all metrics; and a poor-performing, high-risk cluster with negative growth, high volatility, and high downside exposure. Notably, the high-efficiency cluster is concurrent with the top-rated stocks of the composite scoring analysis, which independently proves the ranking methodology.

Previous close versus price stability

The intraday price stability of NIFTY 50 was tested by clustering using a K-means ($K = 8$) model on five bivariate combinations, i.e., Previous Close versus Open, High, Low, Last, and Close prices. Figure 4 illustrates the results of the clustering, with a subplot showing the results of each individual price parameter against the prior close value of the day. Cluster centers (red dots) mark the normal pricing regimes in the dataset. In its assessment, the analysis is intended to identify the extent of consistency between daily price formation and the close of the prior session, hence evaluating intraday continuity and volatility properties.

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

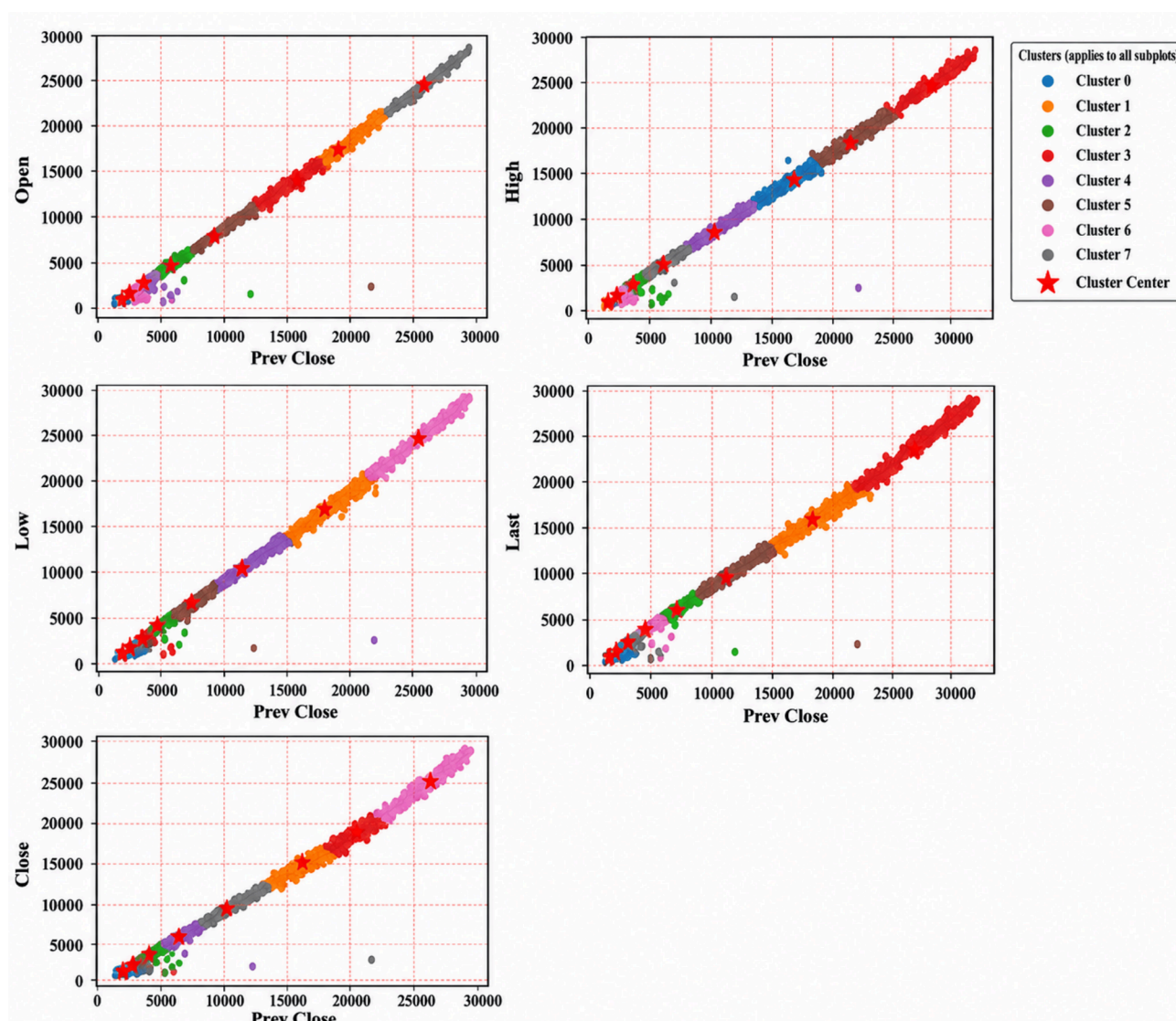


FIGURE 4: Previous Close Versus (Open, High, Low, Last, and Close)

As depicted in Figure 4, there exists a strong linear relationship between the previous close and other price parameters (Open, High, Low, Last, and Close), indicating price continuity in the market. A strong linear correlation along the 45° reference direction can be observed in all five projections; hence, it is evident that the majority of day-to-day prices are strongly tied to the previous day's closing price. The clustering does not reflect directional segmentation but rather price-level segmentation. This indicates a high frequency of intraday fluctuations occurring within a narrow range in terms of absolute price magnitude.

Under the Previous Close and Open projection, the range of cluster centers is within ₹350-360 in the lower band and ₹28,818-28,870 in the highest price band, respectively. The other in-between clusters are the mid-price and index-level regimes, such as ₹1,200-1,202, ₹2,606-2,610, ₹4,780-4,790, ₹8,775-8,786, ₹16,500-16,530, and ₹21,300-21,400. In both cases, there is nearly no difference between the initial price and the concluding value, indicating no or minimal overnight gap formation. The clustering can therefore be termed as price differentiation compared to directional dispersion.

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

The Previous Close and High relationship shows that intraday highs are a little higher than those of the previous close, which shows slight upward excursions of prices on the trading days. At the low prices of stocks and at the very high-priced index-type security, the cluster centers will be ₹366-374 and ₹28,822-29,199, respectively. Mid-range clusters are ₹1,246-₹1,267, ₹2,749-₹2,789, ₹5,270-₹5,345, ₹9,585-₹9,725, ₹16,644-₹16,912, and ₹21,460-₹21,791. These limited vertical changes among the High, Previous Close, and High within the entire clusters denote that volatility is controlled and low, with no high intraday spikes.

The cluster centers of the Previous Close versus Low projection are slightly lower than the previous close, hence indicating a slight intraday downward movement. High-price regime records are ₹27,072-₹26,751, and low-price regime records are ₹332-₹327. Other mid-range clusters are built on nearly ₹1,064-₹1,047, ₹2,240-₹2,210, ₹3,828-₹3,780, ₹6,465-₹6,385, ₹10,693-₹10,536, and ₹18,168-₹17,906. The amount of deviation of the recent close in all the price segments is also fairly low, but in a proportionate way, which proves the action of limited downside shock during the daily sessions.

The Previous Close and Last relationship suggests nearly the same values of the cluster centers for the two variables. As an example, cluster centers are in the ranges of 1,064-1,064, 2,241-2,241, 3,829-3,831, 6,472-6,474, 10,693-10,696, 18,147-18,144, and 27,024-27,029. The fact that the difference between the last traded and the previous closing price is insignificant is a pointer to the fact that most of the securities close the trading vicinities at their initial reference, indicating intraday equilibrium behavior. Similarly, cluster centers have a nearly equal position in every price band in the Previous Close versus Close projection. The groups of low prices are placed around 371-371, while higher prices are 1,265-1,265, 2,783-2,784, 5,333-5,335, 9,664-9,664, 16,674-16,684, 21,503-21,507, and 28,845-28,856. High levels of day-to-day continuity and price anchoring are indicated by the low difference between day-to-day price adjustments and the consecutive closing prices.

Structurally, the results of the clustering show that price size is the primary foundation of segmentation when compared to volatility regime or directional change. We have eight price bands, as we would get in any subplot: low-priced equities, low-priced stocks, high-priced stocks, and price bands at index levels. However, all the bands have negligible intraday changes in the last close. The empirical findings therefore demonstrate that the intraday stability of the universe of the NIFTY 50 is high. Gap-up or gap-down activities are limited, and day-to-day movements of highs and lows are percentage-limited. This type of stability justifies the precision of the calculations of both the returns-based volatility and the risk-adjusted measures, as shown above. In addition, the clustering analysis provides a convenient format for establishing the pricing bands and classification of securities based on absolute price, which can be helpful in supporting portfolio segmentation as well as liquidity planning.

Conclusions

This study designed and tested a comprehensive risk-adjusted analytical framework for evaluating NIFTY 50 stocks by combining financial performance indicators, liquidity screening, K-means clustering, and forecasting using the time-series ARIMA model. In addition to price-based methods, the proposed framework takes into account growth, volatility, downside risk, and market liquidity simultaneously, which makes it a more reliable and practical stock evaluation method. The experimental results show that the composite scoring method is able to identify high-quality stocks with better risk-adjusted performance, and the clustering method can show different market segments of stocks according to their growth and risk factors. Integrating ranking and clustering provides valuable insights into portfolio construction, stock selection, and investment decision-making. In addition, the inclusion of liquidity validation makes the framework more practical and useful by ensuring that the selected stocks are appropriate for trading in the real world. The proposed framework is a comprehensive system that combines quantitative risk assessment with predictive analytics, providing a powerful decision-support tool for investors, portfolio managers, and financial analysts. The application of more sophisticated deep learning models, such as Long Short-Term Memory networks and Transformer-based models, along with the inclusion of portfolio optimization methods in the future, could enhance the accuracy of forecasting and long-term investment performance, thereby extending this work.

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. *Cureus J Comput Sci* 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Santosh S. Shinde, Jagannath E. Nalavade

Acquisition, analysis, or interpretation of data: Santosh S. Shinde, Jagannath E. Nalavade

Drafting of the manuscript: Santosh S. Shinde

Critical review of the manuscript for important intellectual content: Santosh S. Shinde, Jagannath E. Nalavade

Supervision: Santosh S. Shinde, Jagannath E. Nalavade

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue.

Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue.

Conflicts of interest: In compliance with [the ICMJE uniform disclosure form](#), all authors declare the following:

- **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work.
- **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work.
- **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The data (and/or code) supporting this study are openly available in NSE - Nifty 50 Index Minute data (2015 to 2026) at <https://www.kaggle.com/datasets/debashis74017/nifty-50-minute-data>.

Acknowledgements

I, Santosh Suresh Shinde, would like to express my sincere gratitude to my guide, Dr. Jagannath Nalavade, for his invaluable guidance, continuous support, insightful suggestions, and encouragement throughout the course of this research. His expertise and mentorship have been instrumental in shaping this work. I am deeply thankful to the MIT School of Computing, MIT Art, Design and Technology University, Loni Kalbhor, Pune, for providing the academic environment, facilities, and resources necessary to carry out this research successfully. I also extend my heartfelt appreciation to the faculty members and staff of the Department of Computer Science and Engineering for their cooperation and assistance. My sincere thanks go to the providers of the publicly available NIFTY 50 dataset and the open-source research community, whose resources greatly supported this study. Finally, I wish to express my deepest gratitude to my parents, family members, friends, and well-wishers for their constant encouragement, patience, and unwavering support throughout my research journey.

References

1. Aslam B: [Identifying optimistic stocks with K-means clustering algorithm](#). International Review of Economics & Finance. 2025, 104:104579. [10.1016/j.iref.2025.104579](https://doi.org/10.1016/j.iref.2025.104579)

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>

2. Yang S: [Deep reinforcement learning for portfolio management](#). Knowledge-Based Systems. 2023, 278:110905. [10.1016/j.knosys.2023.110905](#)
3. Choi C, Kim J: [Outperforming the tutor: Expert-infused deep reinforcement learning for dynamic portfolio selection of diverse assets](#). Knowledge-Based Systems. 2024, 294:111739. [10.1016/j.knosys.2024.111739](#)
4. Cui T, Du N, Yang X, Ding S: [Multi-period portfolio optimization using a deep reinforcement learning hyper-heuristic approach](#). Technological Forecasting and Social Change. 2024, 198:122944. [10.1016/j.techfore.2023.122944](#)
5. Jang J, Seong N: [Deep reinforcement learning for stock portfolio optimization by connecting with modern portfolio theory](#). Expert Systems with Applications. 2023, 218:119556. [10.1016/j.eswa.2023.119556](#)
6. Zhang B: [Graph attention-based heterogeneous multi-agent deep reinforcement learning for adaptive portfolio optimization](#). Scientific Reports. 2026, 16:2674. [10.1038/s41598-025-32408-w](#)
7. Sun Q, Wei X, Yang X: [GraphSAGE with deep reinforcement learning for financial portfolio optimization](#). Expert Systems with Applications. 2024, 238:122027. [10.1016/j.eswa.2023.122027](#)
8. Grudniewicz J, Ślepaczuk R: [Application of machine learning in algorithmic investment strategies on global stock markets](#). Research in International Business and Finance. 2023, 66:102052. [10.1016/j.ribaf.2023.102052](#)
9. Soleymani F, Paquet E: [Deep graph convolutional reinforcement learning for financial portfolio management - DeepPocket](#). Expert Systems with Applications. 2021, 182:115127. [10.1016/j.eswa.2021.115127](#)
10. Soleymani F, Paquet E: [Financial portfolio optimization with online deep reinforcement learning and restricted stacked autoencoder—DeepBreath](#). Expert Systems with Applications. 2020, 156:113456. [10.1016/j.eswa.2020.113456](#)
11. Jiang Y, Olmo J, Atwi M: [Deep reinforcement learning for portfolio selection](#). Global Finance Journal. 2024, 62:101016. [10.1016/j.gfj.2024.101016](#)
12. de Azevedo Takara L, Santos AAP, Mariani VC, dos Santos Coelho L: [Deep reinforcement learning applied to a sparse-reward trading environment with intraday data](#). Expert Systems with Applications. 2024, 238:121897. [10.1016/j.eswa.2023.121897](#)
13. Shavandi A, Khedmati M: [A multi-agent deep reinforcement learning framework for algorithmic trading in financial markets](#). Expert Systems with Applications. 2022, 208:118124. [10.1016/j.eswa.2022.118124](#)
14. Liu W, Gu Y, Ge Y: [Multi-factor stock trading strategy based on DQN with multi-BiGRU and multi-head ProbSparse self-attention](#). Applied Intelligence. 2024, 54:5417-5440. [10.1007/s10489-024-05463-5](#)
15. Ekmekcioğlu Ö, Pinar MÇ: [Graph neural networks for deep portfolio optimization](#). Neural Computing and Applications. 2023, 35:20663-20674. [10.1007/s00521-023-08862-w](#)
16. Shi Y, Li W, Zhu L, Guo K, Cambria E: [Stock trading rule discovery with double deep Q-network](#). Applied Soft Computing. 2021, 107:107320. [10.1016/j.asoc.2021.107320](#)
17. Cao T, Wan X, Wang H, Yu X, Xu L: [Quantitative stock selection model using graph learning and a spatial-temporal encoder](#). Journal of Theoretical and Applied Electronic Commerce Research. 2024, 19:1756-1775. [10.3390/jtaer19030086](#)
18. Jung NH, Oh T: [Factor-based deep reinforcement learning for asset allocation: Comparative analysis of static and dynamic beta reward designs](#). PLoS One. 2025, 20:e0332779. [10.1371/journal.pone.0332779](#)
19. Wu D, Wang X, Wu S: [Construction of stock portfolios based on k-means clustering of continuous trend features](#). Knowledge-Based Systems. 2022, 252:109358. [10.1016/j.knosys.2022.109358](#)
20. Sezer MO, Gudelek MU, Ozbayoglu AM: [Financial time series forecasting with deep learning: A systematic literature review: 2005-2019](#). Applied Soft Computing. 2020, 90:106181. [10.1016/j.asoc.2020.106181](#)

How to cite this article:

Shinde S S, Nalavade J E (October 25, 2026) Risk-Adjusted Performance Ranking and Clustering Analysis of NIFTY 50 Stocks on the Financial Metrics and Time-Series Forecasting. Cureus J Comput Sci 3 : es44389-026-00336-z. DOI <https://doi.org/10.7759/s44389-026-00336-z>