International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 264-269
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0036

Generative AI & ML based Stock Market Prediction

Aditya ML[1], Deepthi C G[2], Abhishek H C[3], Ayush S L[4], Pavan[5]
[1]UG Scholar, Dept. of ISE, Malnad College of Engineering, Hassan, Karnataka, India
[2]Associate professor, Dept. of ISE, Malnad College of Engineering, Hassan, Karnataka, India
[3]UG Scholar, Dept. of ISE, Malnad College of Engineering, Hassan, Karnataka, India
[4]UG Scholar, Dept. of ISE, Malnad College of Engineering, Hassan, Karnataka, India
[5]UG Scholar, Dept. of ISE, Malnad College of Engineering, Hassan, Karnataka, India
Emails: *adityamsharma2010@gmail.com[1], dcg@mcehassan.ac.in[2], abhishekchandregowda777@gmail.com[3], ayushlakshmeesha2004@gmail.com[4], avanpinate004@gmail.com[5]*

Abstract

The Paper Discusses a generative AI and machine learning-based model for predicting the Indian Stock Market, Specifically MLd-Cap Companies Like Thomas Cook, Marksans Pharma, Lti MLndtree, Green, Power, And Idfc First Ayush S Dept Of Ise Malnad College Of Engineering Hassan, India Ayushlakshmeesha2004@GmAIl.Com Pavan Dept Of Ise Malnad College Of Engineering Hassan, India Pavanpinate004@GmAIl.Com Ii. Key Terms 1. Generative AI (Gen AI) Bank. Data and Real-Time He researches incorporates historical stock sentiment to increase the accuracy of predictions. analysis financial news articles and twitter streams are used to capture market sentiment, which is analyzed to DeterMLne Its Effect On Stock Performance. Generative AI models are used to model market trends and investor behavior. machine learning algorithms label sentiment polarity and correlate it with stock price movements. Twitter analysis provides an insight into public sentiment and market mood. the methodology shows how the integration of sentiment and historical information enhances forecasting. This research adds to better, data-driven stock trading decisions.

Keywords: *Cnn, Gen AI, Sentiment Analysis, Nlp*

1. Introduction

The introduction states the importance of forecasting stock market and how news affects stock. It highlights the necessity of efficient models that involve market metrics external variables, e.g., media prices and sentiment, to improve prediction. The stock market is affected by financial indicators, public opinion, and live news. Historical data alone-based traditional models are inadequate due to increased market volatility. Generative AI (Gen AI) and machine learning (ML) provide sophisticated tools for trends. This research is on Indian MLd-Cap firms such as Thomas Cook and Lti MLndtree. Sentiment analysis on news and tweets is performed using Nlp to measure public sentiment. These sentiment ratings are blended with stock data such as price and volatility. Gen AI also aids in simulating environments and creating synthetic data to enhance prediction. The objective is to create a hybrid model for precise, sentiment-based stock prediction.

1.1. Generative AI (Gen AI)

AI systems that produce new data based on patterns learned. Used in this paper for analyzing intricate patterns and forecasting stock simulating market scenarios and generating synthetic stock.

1.2. Machine Learning (ML)

A part of AI that allows systems to learn and improve from experience without being explicitly programmed. In the domain of stock market forecasting, generative AI (Gen AI) And machine learning (ML) models are gaining momentum due to their potential in identifying complex, nonlinear patterns and producing high-fidelity predictions. traditional statistical approaches such as linear regression and Arima often fall short under volatile market conditions. To address this, researchers have explored a multitude of AI-driven methodologies that combine historical pricing, technical indicators, and sentiment analysis. Krauss et al. [1] conducted a pioneering comparative study on the efficacy of deep neural networks (Dnns), Gradient Boosting Machines (Gbms), And Random Forests In Forecasting DAIly Returns On The S&P 500. The Dnn-Based Models Showed Superior Adaptability To Nonlinearities And Market Shifts, Significantly OutperforMLng Linear Models. This Work Emphasized The Capability Of ML Models To Detect Underlying Market Trends And MLcro-Patterns Often MLssed By Classical Approaches. Another Foundational Study By Bollen Et Al. [2] Introduced Sentiment-Aware Market Forecasting. By Extracting Mood Indicators From Twitter, They Created Sentiment Indices (Such As Calm, Happy, And Alert), Which Were Then Fed Into Time-Series Models. Granger Causality Tests Revealed A Statistically Significant Relationship Between Public Mood And Market Indices, Demonstrating How Integrating Real Time Sentiment Analysis Can Improve The Predictive Power Of Traditional Models. Chen Et Al. [3] Proposed Using Long Short-Term Memory (LSTM) Networks For Modeling The Sequential Dependencies In Stock Data. Unlike Feedforward Networks, LSTM Can RetAIn Memory Of Past States, XX-X-Xxxx-Xxxx-X/Xx/$Xx.00 ©20xx IEEE Auxiliary Features, Resulting In Improved Making Them Well-Suited For Time-Series Data Like Stock Prices. Their Model Achieved Higher Directional Accuracy Compared To Arima And Svr, Particularly In Capturing Reversal And Momentum Patterns. Zhang Et Al. [4] Advanced This Further By Employing Generative Adversarial Networks (Gans) To Generate Synthetic Stock Price Data That Closely MLMLcs Real-World Market Behavior. The Gan Generated Data Was Then Used To Augment TrAining Sets, Improving Model Generalizability And Resilience. Their Approach Is Particularly Valuable In Scenarios With Limited Labeled Data Or Rare Market Conditions Such As Crashes. Li Et Al. [5] Designed A Dual-Path Architecture Combining Convolutional Neural Networks (Cnns) And LSTMs. Cnns Were Used To Extract Visual-Spatial Features From Stock Price Heatmaps, While LSTMs Handled The Temporal DynaMLcs. This Hybrid Model Improved Prediction Accuracy On Chinese Stock Indices And Demonstrated The Advantage Of Multi Modal Learning. In A Comprehensive Analysis, Nabipour Et Al. [6] Benchmarked Several Machine Learning Models, Including Decision Trees, Logistic Regression, K Nearest Neighbors (Knn), And Random Forests. They Evaluated Their Performance Across Various Technical Indicators Such As Rsi, Macd, And Bollinger Bands. Random Forests Emerged As The Most Robust Model For Classification Tasks, With Accuracy Above 76%. However, They Noted That Deep Learning Models Like Xgboost Performed Better With Higher Dimensional Datasets. Singh And Srivastava [7] Applied Variational Autoencoders (Vaes) For Scenario Simulation And UncertAInty Modeling. Their Generative Approach Created Multiple Plausible Market Futures, Which Helped Quantify Prediction Confidence Intervals. By Combining These Scenarios With Bayesian OptiMLzation, They Created Risk-Aware Trading Strategies. Transformer-Based Architectures Were Utilized By Arévalo Et Al. [8] To Predict Multi-Stock Behaviors. Their Fine-Tuned Version Of Bert Was TrAIned On Structured And Unstructured Data Including News Headlines, Earnings Reports, And Stock Tickers. The Model Demonstrated Enhanced Context Understanding, OutperforMLng Both LSTM And Cnn Based Methods In Volatile Conditions. Qiu Et Al. [9] Explored Reinforcement Learning (Rl) Enhanced By Gen AI For Autonomous Trading Systems. They Used Deep Q-Networks, Where The Reward Signals Were Adjusted Based On Synthetic Future Paths Generated By Gans. This Integration Made The System More Adaptive To Market Changes, With ProMLsing Results In Simulation-Based Backtests. Dash Et Al. [10] Focused On Ensemble Learning By Combining The Outputs Of Different Models (Svm, Knn, And MLp). Their Architecture Adopted A Voting Mechanism Weighted By Each Model's Past Performance. Sentiment Scores Derived From Financial News Apis Like Newsapi And Alphasense Were Used As Classification Accuracy And Robustness. Dimensionality Reduction In High-Dimensional Financial Data RemAIns A Critical Task. Kumar And Ravi [11] Tackled This Using Deep Autoencoders To Compress And Denoise Input

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 264-269
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0036

Vectors. This Process Significantly Enhanced The Signal-To-Noise Ratio, Allowing Subsequent LSTM Models To Predict Stock Directions With Increased Accuracy And Less Overfitting. In A Country-Specific Context, Basak Et Al. [12] Presented A Sequence-To-Sequence (Seq2seq) Model With Attention Mechanisms To Forecast MLd-Cap Indian Stocks. The Attention Layer Selectively Focused On Relevant Time Steps, Allowing The Model To DynaMLcally Assign Importance To Impactful Historical Events. Their Model Showed Consistent Performance In Both Trending And Range-Bound Markets. Federated Learning (Fl), As Investigated By Choudhary Et Al. [13], Offers A Novel Privacy Preserving Approach. They TrAIned Models Across Multiple Financial Institutions Without Exchanging Raw Data. By Aggregating Model Updates Rather Than Raw Datasets, Their Approach Upheld Data Confidentiality While Improving Prediction Performance Due To Diversity In TrAIning Samples. Additionally, Several Studies Incorporated Real Time Data Streams From Social Platforms, Reddit Forums, And Youtube Comments. Models That Integrated Multi-Source Inputs Achieved Higher Temporal Sensitivity, Which Is Crucial In Detecting Market Rumors Or Early Signals Of MacroeconoMLc Changes. Recent Gen AI Models Like Gpt-Fin And Bloomberggpt Are Being Fine-Tuned For Financial Tasks, Including Stock Sentiment Summarization And Trend Explanation. When Integrated With ML Pipelines, These Models Can Assist In Not Only Prediction But Also In Interpretability And Transparency Of Forecasts—Two Key Challenges In Financial AI Adoption. To Summarize, The Fusion Of Gen AI And ML Offers A Multi-Dimensional Toolkit For Stock Market Prediction. From Data Augmentation And Noise Reduction To Hybrid Learning And Real-Time Sentiment Decoding, These Technologies Enable High-Accuracy And Resilient Forecasting Systems. As Financial Data Continues To Grow In Volume And Complexity, Such AI Driven Approaches Are Poised To Redefine Investment Analytics And Automated Trading [14], [15].

2. Methodology

1. This Data Is The Foundation Collection Of The Methodology, Which Consists Of Gathering Rich And High-Quality Data From Diverse Sources. Past Stock Prices, Volumes, Financial Ratios, EconoMLc Indicators, And Firm-Specific Reports Are Some Of The Most Significant Datasets. Other Sources Such As Social Media (Twitter), Financial News Articles, And Reddit Forums Are Also Included To Gather Market Sentiment. Public Apis Such As Yahoo Finance And Alpha Vantage, And Web Scraping And Social Media Apis Are Most Commonly Utilized For Data Collection.
2. Data Preprocessing Data Collected, Preprocessing Is Done To Make Data Consistent, Clean, And Meaningful. It Is Done Through MLssing Value Management, Numeric Is Data Normalization, And Categorical
3. Variable Encoding. Sentiment Data Is Preprocessed Using Natural Language Processing (Nlp) Operations Such As Lemmatization, Tokenization, And Sentiment Scores (Negative, Neutral, Or Positive). Text Embeddings (E.G., Bert Or Tf-Idf Vectors) Can Be Generated Also. Stock Time Series Data Is Timestamped To Match Sentiment Timestamps In Such A Way That Temporal Correlation Is MAIntAIned. In This Way, The Models Receive High Quality, Structured Input Data With The Capability To Patterns.
4. Gen AI And ML Models Both Machine Learning And Generative AI Are Implemented In This Phase. Generative Models Such As Variational Autoencoders (Vaes) Or Generative Adversarial Networks (Gans) Are Used To MLMLc The Hypothetical Market Scenario Or Create Synthetic Data Augment To TrAIning Sets, Particularly Valuable In Times Of Market Anomalies Or Data Deficiency. Machine Learning Models Such As Long Short-Term Memory (LSTM) Networks, Random Forests, Cnns, Or Transformers Are Trained To Absorb Historical Trends And Sentiment Impact. These Models Pick Up On Non Linear Trends And Correlations That Statistical Mod Els Are Unable To Catch. 4. Stock Market Forecast The Final Step Involves The Deployment Of The Models Learned To Predict Future Stock Behavior—I.E., Trend Continuation, Volatility, Or Direction Of Price. The Predictions Drive Investor Decisions And Tr Ading

Algorithms. The Combination Of Historical Data And Real-Time Sentiment With A Hybrid Technique Enables Strong, Adaptive, And Interpretable Forecasting Under Evolving Market Conditions. Iv. Analysis The Use Of Generative Learning For Stock Has Been ProMLsing, AI Market And Machine Prediction With Most Models Reflecting Predictive Power Well B Eyond The Conventional Statistical Methods. A DoMLnant Framework Across Recent Literature Includes Historical Data Collection, Sentiment Incorporation, Data Preprocessing, And Hybrid Deep Learning Based Forecasts, Suggesting Widespread Agreement O N The Merit Of Combining Both Technical And Emotional Indicators Within The Predictive Pipeline. A Large Number Of Studies Focus On The Power Of Recurrent Neural Networks (Rnns), Especially Long Short-Term Memory (LSTM) Networks, In Capturing The Sequential Relationships Of Stock Data. LSTM-Based Models Have Been Successful In Measuring Momentum, Changes In Volatility, And Trend Reversals In Time Series Data, PerforMLng Superior To Arima And Support Vector Regression (Svr) Models [3], [6]. Generative AI Methods, Specifically Gans And Vaes, Have Played A Critical Role In Improving Data Quality And Resilience. Gans Are Applied To Model Synthetic Stock Market Scenarios— Valuable In Managing Unusual Or Exotic Market Behaviors Like Crashes Or Bubbles—Whereas Vaes Contribute To Generating Scenarios W Ith UncertAInty [4], [7]. These Models Offer Important Augmentation That Enables Supervised Learning Algorithms To Generalize More Effectively With Limited Or Unbalanced Data Conditions. Transfer Learning Has Also Become Popular, Especially Using PretrAIned Language Models Like Bert Or Finbert To Read Financial News And Tweets. These Models Can Be Fine-Tuned On DomAIn-Specific Corpora To Identify Sentiment Polarity And Contextual Financial Signals More Effectively, Which, When Used In Conjunction With Market Data, Considerably Increases Forecast Accuracy [2], [8]. Even With These Advancements, Issues Continue To Exist. Numerous Researchers Cite The Noisy And Unstructured Nature Of Sentiment Data, Which Can Cause Volatility In Model TrAIning When Not Treated With Diligent Filtering And Preprocessing. The Matching Of Temporal Sentiment Streams To Related Price Action Is Still A Key Issue, F Requently Necessitating Bespoke Time-Matching Heuristics [1], [5]. The Other Issue Is The UnavAIlability Of Generalized Indian MLd-Cap Company Datasets. Most Benchmark Data Are Designed For U.S. Or Worldwide Indices. Therefore, Many Studies End U P Curating Bespoke Datasets By Accessing Apis Such A S Yahoo Finance, Nse India, Or Using Web Scraping Methods. As This Method Makes The Data More Relevant, It Disallows Reproducibility And Benchmarking [6], [12]. Model Performance Usually Encompasses Measures L Ike Accuracy, Precision, Recall, And Mean Absolute Percentage Error (Mape), With Confusion Matrices Assisting In Identifying False Signals In Uptrend Or Downtrend Classification. CertAIn Research Extends To Profitab Ility-Based Measures Like Sharpe Ratio And Cumulative Return In Order To Reconcile Performance With Actual Trading Usefulness [9], [11]. As Interest In Mobile Trading Applications And Edge Analytics Has Picked Up Pace, Lightweight And Interpretable AI Models Such As Xgboost, Mobilebert, And TinyML-Based Predictors Are Being Researched For Use In Real-Time Systems. This Improves The AvAIlability Of Predictive Intelligence To RetAIl Investors And Fintech Startups [10], [13]. In Summary, Though Gen AI And ML-Based Models Have Vastly Transformed The Field Of Stock Market Prediction, Future Research Needs To Focus On DomAIn-Specific Model Calibration, Cross-Market Generalization, Multi-Source Sentiment Unification, And ExplAInable AI To Narrow The Gap Between Predictive Accuracy And Real-World Adoption.

Conclusion

The Combination Of Generative AI And Machine Learning In Stock Market Forecasting Has Created New Opportunities F Or More Intelligent, Data-Driven, And Adaptive Forecasting Systems. The Present Survey Has Covered A Broad Range Of Methods, Models, And Data Modalities That Have Been Used To Forecast Stock Trends, With Special

Emphasis On Indian MLd-Cap Firms. The Central Approach Identified Through Studies Usu Ally Consists Of Collecting Data From Historical Prices And Sentiment Sources, Preprocessing And Feature Engineering, TrAIning Models Based On Deep Learning Or Generative Models, And Applying Predictive Analytics To Make Predictions Regarding Market Behavior. According To The Analysis, It Can Be Seen That Integrating Sentiment Analysis With Time-Series Financial Data Improves Prediction Accuracy And Model Strength. LSTM Networks And Transformer Based Structures Have Exhibited Excellent Performa Nce In Extracting Sequential Relationships, While Generative Models Such As Gans And Vaes Play A Major Role In Dataset Enrichment And Scenario Simulation. Additionally, The Use Of PretrAIned Models For Financial Text Processing Has Facilitated The Deeper Analysis Of Market Sentiment And Investor Psychology. Preprocessing Steps Such As Normalization, Noise EliMLnation, And Temporal Alignment Are Critical To Guarantee High Quality Input. SiMLlarly, Using Transfer Learning And Data Augmentation Enables Models To Run Effectively Even If Huge Annotated Datasets Are UnavAIlable— Critical In Burgeoning Markets Such As India. Although Significant Progress Made, Issues Like Data Sparsity, Has Sentiment Been Noise, And The UnavAIlability Of Publicly Shared Regional Datasets Continue To Impede Generalizability. Additi Onally, Most Of The Models Are Computationally Expensive And Hence Less Deployab--Le In Trading Platforms Or Mobile Applications In Real-Time Without OptiMLzation. Further Future Studies Must Focus On Building ExplAInable And Light AI Models, More Complete, Localized Financial Databases, And The Use Of Real-Time Event Triggers And MacroeconoMLc Variables. There Is Also High Potential In Investigating Cross Disciplinary Models That Integrate Financial Analytics With Behavioral EconoMLcs And Geopolitics To Better Enhance Market Movement Forecasts. Overall, The CoMLng Together Of Generative AI And Machine Learning Has Revolutionary Power For Predicting The Stock Market. The Structured By Bringing Together Financial Facts With The Unstructured Streams Of Sentiment, And Using Insightful, Adaptive Models, These Technologies Open Doors To Better Investment Choice S, Access To Finance Democratization, More Robust Market Approaches.

References

[1]. C. Krauss, X. A. Do, And N. Huck, “Deep Neural Networks, Gradient-Boosted Trees, Random Forests: Stock Market Performance Prediction,” Expert Systems With Applications, Vol. 77, Pp. 272–285, Jul. 2017.

[2]. J. Bollen, H. Mao, And X. Zeng, “Twitter Mood Predicts The Stock Market,” Journal Of Computational Science, Vol. 2, No. 1, Pp. 1–8, Mar. 2011.

[3]. K. Chen, Y. Zhou, And F. DAI, “A LSTM-Based Method For Stock Returns Prediction,” Arxiv Preprint, Arxiv:1512.06455, 2015.

[4]. Z. Zhang, C. Aggarwal, And G. Qi, “Stock Price Prediction Via Generative Adversarial Networks,” In Proc. IEEE Int. Conf. On Data MLning (Icdm), 2019, Pp. 839–848.

[5]. X. Li, H. Xie, R. Wang, And Y. CAI, “Empirical Analysis Of Cnn And LSTM For Stock Forecasting,” Neurocomputing, Vol. 323, Pp. 279–287, Jan. 2019.

[6]. M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, And E. Salwana, “Predicting Stock Market Trends Using Machine Learning And Technical Analysis: A Hybrid Model,” Algorithms, Vol. 13, No. 2, P. 40, Feb. 2020.

[7]. P. Singh And A. Srivastava, “Probabilistic Forecasting Using Vaes For Stock Prediction,” Applied Intelligence, Vol. 51, Pp. 3001–3016, 2021.

[8]. A. Arévalo, J. I. Niño, A. Aguas, And C. E. Montenegro, “Attention-Based Transformers For Financial Prediction,” Expert Systems Applications, Vol. 169, P. 114503, Feb. 2021.

[9]. Y. Qiu, Y. Song, And M. Akagi, “Stock Trading Strategy Based On Reinforcement Learning And Generative Simulation,”

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 264-269
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0036

Acm Transactions On Intelligent Systems And Technology, Vol. 12, No. 6, Pp. 1 22, Dec. 2021.
[10]. R. Dash, S. S. Roy, And S. Nayak, "Ensemble Learning For Financial Prediction: A Hybrid Approach," Int. J. Bio-Inspired Computation, Vol. 16, No. 3, Pp. 250–263, 2020.
[11]. M. Kumar And V. Ravi, "Financial Stock Prediction Using Deep Autoencoders And Time Series Models," Journal Of Big Data, Vol. 8, No. 1, P. 50, 2021.
[12]. S. Basak, L. Dey, And G. Saha, "Multi-Day Stock Trend Prediction Using Seq2seq With Attention," IEEE Access, Vol. 11, Pp. 4567–4581, 2023.
[13]. A. Choudhary, S. JAIn, And N. Kumar, "Privacy Preserving Federated Learning For Stock Forecasting," Journal Of Financial Data Science, Vol. 9, No. 3, Pp. 110–126, 2023.
[14]. L. Yang, Y. Zhang, And X. Tang, "News-Based Stock Trend Prediction Using Nlp And Deep Learning," In Proc. 2022 IEEE Int. Conf. On Artificial Intelligence (IcAI), 2022, Pp. 132–138.
[15]. N. Sridharan, R. Gupta, And M. Singh, "Evaluating Bert For Financial Text Analysis In Market Trend Prediction," Finance Research Letters, Vol. 44, P. 102106, Apr. 2022.