Integrating Heterogeneous Data for Stock Market Prediction: A Systematic Literature Review

Abdullah Almusned, Mohammad Mehedi Hassan, Bader Alkhamees, Muhammad Al-Qurishi
Department of Information System-College of Computer and Information Sciences, King Saud University, Saudi Arabia

***Abstract*—This systematic literature review examines recent developments in stock market prediction using heterogeneous data sources that combine technical indicators, fundamental attributes, and sentiment-driven signals. Despite the growing adoption of machine learning in financial forecasting, existing research remains fragmented across data modalities, fusion strategies, and evaluation protocols, limiting comparability and practical applicability. Studies published between 2018 and 2024 were retrieved from five major scholarly databases and screened based on predefined eligibility criteria, resulting in 44 peer-reviewed articles included in the final analysis. The review synthesizes the quantitative and qualitative data modalities employed, the machine learning and deep learning methodologies adopted, the evaluation metrics used to assess predictive performance, and the principal challenges associated with multi-source stock market prediction. Findings reveal a clear shift toward deep learning architectures, hybrid fusion techniques, and the integration of external information such as news, corporate disclosures, and social media sentiment. Despite this progress, the literature exhibits inconsistent evaluation practices, limited attention to temporal data leakage, and insufficient coverage of non-English and emerging markets. This review consolidates current knowledge, presents a structured taxonomy of heterogeneous data sources and fusion strategies, and identifies open research challenges to guide future work in multimodal stock market prediction.**

Keywords*—*Stock prediction; heterogeneous data; machine learning; quantitative and qualitative data; systematic review

I. Introduction

Stocks are one of the most profitable investment sectors in the financial market. A company's stock can be defined as "the stock or share or equity is a financial instrument representing that the holder has proportionate ownership of the issuing an organization's assets and its profit" [1]. The companies issue those stocks and sell them to investors by the owners to raise funding for further development of the company or to solve the problem of temporary capital scarcity. They sold the initial stocks through an Initial Public Offering (IPO) before they got listed in the market for the first time. The stock market is the place where investors can buy or sell their stocks (shares) with each other. There are many financial products for investors, and stocks are one of the most easily available products [2]. They are trying to buy the stock when the price drops and sell it when it increases.

Random walk theory assumes that it is impossible to predict the stock based on only historical stock prices [2]. Stock price fluctuation depends on many factors and can be divided into two groups: Technical indicators (TI), which predict the stock by statistical equations based on historical data, and Fundamental indicators (FI), which focus on company performance and standing [3]. Furthermore, the Efficient Market Hypothesis (EMH) semi-strong form states that prices are adjusted frequently according to market and public information [4], [5]. Since much information needs to be processed before predicting the stock future, human brains are incapable of doing it. Machine learning and data mining were used to overcome this problem, find correlations, and predict the future based on all available information, such as historical stock price (HSP), social media, news, and furthermore.

Several review papers have examined stock market prediction; however, the majority focus exclusively on a single data source, typically historical price movements, and evaluate how past prices alone can forecast future trends [6], [7], [8], [9]. While some reviews have considered the use of heterogeneous data, they often do so without explicitly integrating both fundamental and technical analyses into a unified predictive framework [10], [11], [12], [13]. Other studies have focused specifically on heterogeneous data expressed in textual form or have narrowed their scope to examining how news-based information can be incorporated into prediction models [14], [15]. Furthermore, one review concentrated exclusively on the U.S. market, overlooking the structural and behavioral differences present in other global markets [16].

A notable study in this domain is the 2020 systematic literature review titled "A systematic review of fundamental and technical analysis of stock market predictions", which covered research published between 2008 and 2018 [17]. Out of 122 analyzed studies, only 13 incorporated both fundamental and technical indicators. Building upon the endpoint of that review, our SLR extends this line of inquiry by examining more recent contributions.

The present systematic literature review is guided by four primary objectives. First, to identify and classify the distinct types of data sources employed in stock market prediction research published between 2018 and 2024, including both quantitative and qualitative modalities. Second, to examine and compare the machine-learning and deep-learning methodologies adopted across the reviewed studies, with particular attention to fusion strategies and architectural design. Third, to survey the evaluation metrics and validation protocols used to assess predictive performance, and to assess their suitability for financial time-series tasks. Fourth, to identify the principal challenges and open research problems associated with multi-source stock market prediction, and to derive evidence-based recommendations for future work.

This review makes several contributions to the growing body of literature on AI-driven financial forecasting. By systematically synthesizing 44 studies spanning seven years, this study provides a timely and comprehensive update to prior surveys that either predate the deep-learning era or focus exclusively on a single data modality. More importantly, it addresses an analytical gap in the existing literature by evaluating heterogeneous data integration not merely as a feature-engineering exercise, but as a methodological design challenge involving temporal alignment, fusion strategy selection, and leakage-safe validation. The findings are expected to benefit researchers seeking to build on existing work, practitioners designing data pipelines for algorithmic trading systems, and reviewers evaluating the methodological rigor of future submissions in this domain.

This systematic literature review synthesizes heterogeneous-data approaches for stock-market prediction and provides:

- Comparative synthesis: a structured, evidence-based comparison linking modality choice, temporal alignment, fusion strategy, and leakage-safe evaluation across 44 reviewed studies, consolidating findings that were previously scattered across disconnected works.
- Modality characterization: a formal taxonomy of how market/technical, fundamental, and text-based sentiment sources are employed across different forecasting horizons, providing a reference framework for future study design.
- Method and fusion comparison: a systematic comparison of dominant model families — including classical machine learning, deep sequence models, and reinforcement learning frameworks — alongside early, late, and hybrid fusion architectures, with an explicit assessment of their practical strengths and limitations.
- Evaluation and guidance: a summary of common evaluation practices across the reviewed literature, accompanied by concrete recommendations to improve rigor, reproducibility, and cross-study comparability in future stock market prediction research.

The remainder of this study is organized as follows. Section II presents the methodology. Sections III and IV cover the execution stage and results, respectively. Section V presents the discussion, and Section VI provides the conclusion

II. Methodologies

This section conducts a Systematic Literature Review (SLR) to explore the literature. SLR is a systematic technique that combines all the latest research publications on a particular topic or research questions [18]. This reference offers a framework that was followed in conducting this review. The SLR is conducted systematically on the stock market prediction with technical and fundamental analysis.

SLR methodology consists of multiple steps starting from specifying the research questions about the problem and then conducting the review by identifying the search, applying inclusion/exclusion criteria, a manual search on defined libraries, and finally filtering based on the quality assessment, as illustrated in Fig. 1. After collecting the publications, start Data extraction and monitoring, then data synthesis. Finally, reporting the findings. The review is divided into three stages as [19] proposed: the preparatory stage, the execution stage, and the presentation stage.

A. Research Questions

To determine the scope of the SLR, the following research questions are specified:

RQ1: What are the distinct types of data sources used for stock market prediction?

RQ2: What methodologies are used to predict stock behavior?

RQ3: What are the different performance parameters used in stock market prediction?

RQ4: What are the main difficulties in predicting stock market behavior through multi-source approaches?

B. Search String and Databases

The search strategy is identified as follows. The internet is full of digital libraries, which contain millions of publications. Five top libraries suitable for the field of machine learning were selected. These libraries are expected to encompass all significant research in this area. The libraries are IEEE, Springer, Wiley, ScienceDirect (Elsevier), and Emerald. The search was based on keywords. The keywords were joined by applying the “AND” and “OR” operators to cover the most publications related to the research questions. This search string maximizes retrieval coverage across each library in stock prediction and machine learning, technical and fundamental analysis.

The following search string was used:

“("stock*") AND ("predict*" OR "forecast*") AND ("machine learning" OR "deep learning") AND ("technical analysis " OR " quantitative analysis") AND ("fundamental analysis " OR "qualitative analysis" OR "sentiment analysis")”

TABLE I. Inclusion and Exclusion Criteria

Inclusion	Exclusion
Use of English as the primary language	studies those are not relevant to the research questions
Period of publication between 2018 and 2024	Articles outside the domain of ML
The primary purpose is to predict the stock market	Papers that did not focus on aspects of Stock Prediction
Articles found in complete form	The study reports without evaluation metrics
Available from the sources selected for the review	Publication type other than a journal

C. Inclusion and Exclusion Criteria

Specific criteria must be established in advance to ensure a fair representation of the research publications and eliminate irrelevant publications. It should be noted that the primary goal of this SLR was to shed light on recent stock predictions by using fundamental and technical analysis difficulties. The search was limited to 2018-2024, which is after [17] SLR and till the end of last year. Moreover, the only publication that will be accepted is one written in English. Table I contains the complete list of inclusion and exclusion criteria. While screening the title and abstract, adherence to the inclusion and exclusion criteria was verified.

D. Quality Metrics

After downloading the nominated studies, the quality and accuracy of the data given in the research were assessed after completing a thorough study of the reviewed material. This was achieved using the following questions:

QA1: Is there a clear statement of the aims and objectives of the research?

QA2: Is the reviewed study's central theme connected to Stock Prediction?

QA3: Do they use two or more sources of data?

QA4: Is there an adequate description and justification of the proposed contribution, method, or approach?

QA5: Were the research findings stated?

QA6: Are evaluation procedures well described and investigated?

QA7: Is the study transparent and objective regarding the data collection procedures adopted? Furthermore, was the data sufficient?

QA8: Do the expressions and reporting of the results support the findings of the research?

With the requirements and criteria defined, the preparation stage is complete.

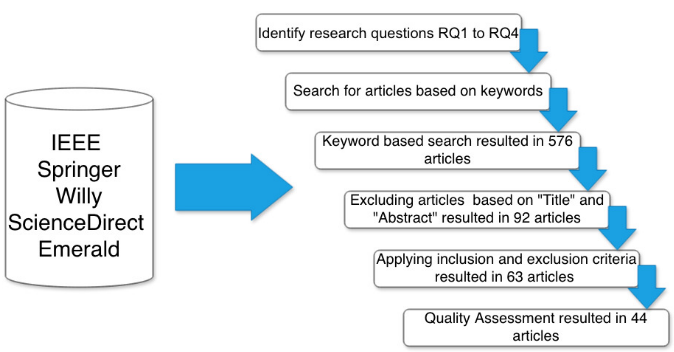

Fig. 1. SLR filtration process.

III. Execution Stage

The previously mentioned libraries were used to conduct the search with the search string specified earlier, yielding 576 studies. Afterwards, the results were screened by reading titles and abstracts and filtering unrelated publications based on the inclusion and exclusion criteria. This process yielded 63 studies requiring full reading and quality assessment. Finally, 44 studies answered the research questions and satisfied the eligibility criteria. Fig. 1 summarizes the steps. Those 44 studies combine technical and fundamental analysis. A notable increase in research interest is observed: whereas [17] identified only 13 studies over ten years, the present review identified 44 in the following seven years, as shown in Fig. 2. At the end of this stage, data were extracted from the studies and prepared for synthesis.

Table II presents a comprehensive summary of the reviewed studies, indicating the markets in which the companies operate, the additional data sources used alongside daily stock prices, the duration of data collection in years, and the total number of indices or companies included in each study.

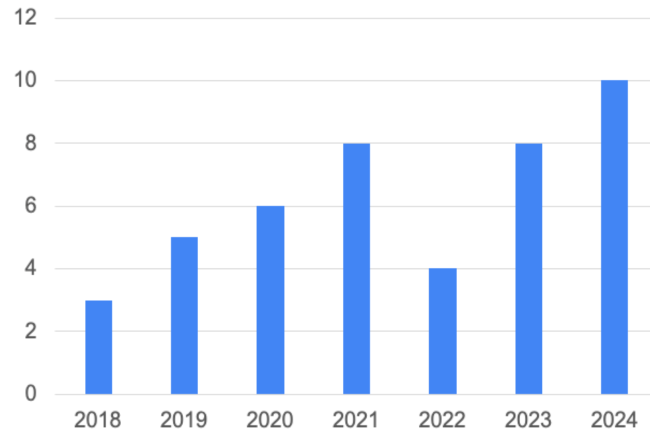

Fig. 2. Reviews per year.

IV. Results

This section addresses the SLR research questions based on the analyzed studies.

A. RQ1: What are the Distinct Types of Data Sources used for Stock Market Prediction?

The recent literature on stock market prediction reflects a systematic effort to integrate heterogeneous data sources, which can be broadly classified into quantitative and qualitative categories. Quantitative data consists of structured numerical information suitable for statistical analysis and model computation. Historical stock prices (HSP)—including open, close, high, low, and trading volume for each trading day—form the core of this category, serving both as direct model inputs and as the basis for calculating technical indicators to analyze market trends. Company reports, such as annual financial statements and quarterly earnings disclosures, provide structured measures of financial health that are central to fundamental analysis. Additional quantitative streams include search engine trend indices (e.g., Google Trends, Baidu Index), which capture public attention and information-seeking behaviors, as well as macroeconomic indicators and industry-specific metrics that reflect broader economic forces. By contrast, qualitative data sources primarily consist of unstructured or semi-structured text, such as financial and public news (including comment sections), which encapsulate investor sentiment and market reactions. Other qualitative channels include collaborative platforms such as Wikipedia, social media networks (e.g., X/Twitter, Sahamyab), and investor forums where opinions and forecasts are exchanged. A growing number of studies employ hybrid approaches that combine quantitative and qualitative inputs, aiming to leverage the complementary strengths of each data type.

Fig. 3 shows a taxonomy that groups the different qualitative and quantitative data and the type of each category.

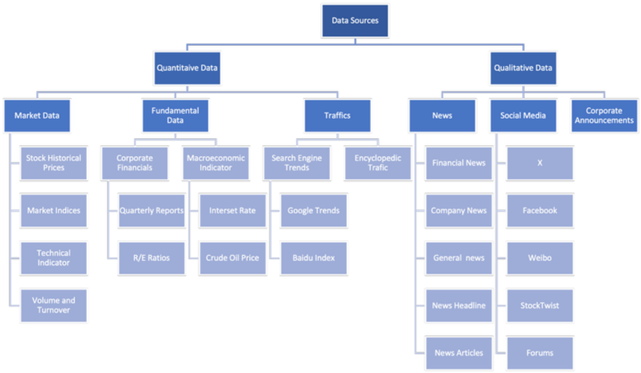

Fig. 3. Taxonomical classification for the data used for building the models.

Empirical investigations span a diverse range of markets, underscoring the global applicability of multimodal prediction frameworks as shown in Fig. 5. Approximately 70% of studies focus on large and information-rich environments such as the United States and China, where data availability and media coverage support both high-frequency and long-horizon modeling. Japan and Iran illustrate contexts in which local news sentiment and domestic factors significantly influence valuation. Further studies encompass India, Pakistan, Hong Kong, Brazil, and Turkey, extending multi-source fusion methodologies into emerging market settings. Other markets such as Australia, the UK, Indonesia, Poland, and Ghana contribute additional perspectives, each shaped by distinct regulatory frameworks, investor behaviors, and macroeconomic conditions .

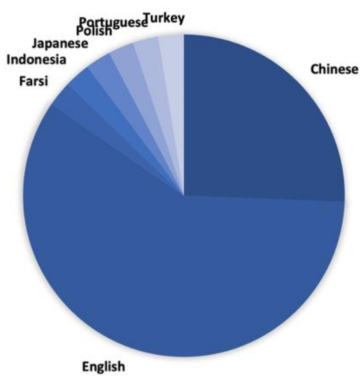

Fig. 4. Language used in studies.

Regarding the linguistic dimension of the utilized data sources, the reviewed studies demonstrate that stock prediction is inherently multilingual, as textual sentiment and information must be processed in the languages most relevant to each market, as shown in Fig. 4. English remains the dominant language, given its prevalence in international financial news, multinational corporate disclosures, and widely used social media platforms. English has not been used only for US stock market data; it has also been used for Indonesia, Pakistan, India, and Hong Kong. However, Chinese is increasingly important for analyses focusing on Mainland China and Hong Kong, particularly when leveraging local news outlets and social media. Japanese language sources are critical for predictions related to the Tokyo Stock Exchange, while Farsi is used for Iran's domestic market data. Furthermore, Portuguese appears in studies analyzing the Brazilian market, Polish is employed for Poland's stock forum data, and Turkish supports sentiment extraction and trend prediction in Turkey's financial sector. The linguistic diversity of financial texts requires advanced Natural Language Processing (NLP) methods capable of addressing domain-specific terminology, local linguistic variations, and cross-language sentiment lexicons.

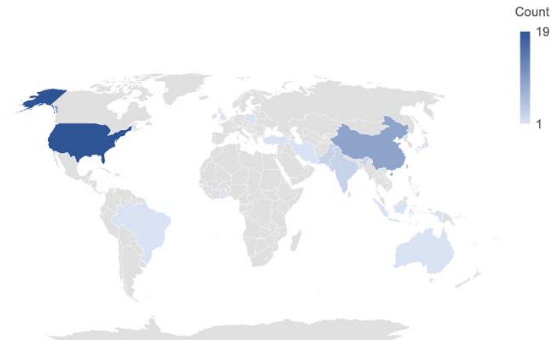

Fig. 5. Markets used in the studies.

The scope of companies examined in prior studies varies substantially, reflecting differences in research objectives, data accessibility, and forecasting aims. Several works concentrated on a single market index, such as the S&P 500, or a sector-specific index, such as the DJIA, treating these benchmarks as proxies for broader market dynamics. Others targeted a small number of companies, often ranging from 2 to 5 firms, which are suitable for fine-grained analysis or case-specific modeling. A moderate number of studies included company sets between 10 and 40, with some explicitly selecting firms from different sectors. A few studies adopted a hybrid approach by combining a group of companies with one or more indices—for example, 8 companies in addition to 3 indices and 5 companies with 1 index—to capture both firm-level and aggregate market signals. Larger-scale investigations included 88, 180, and even 240 companies, aiming for broader generalizability. Additionally, this heterogeneity highlights the flexibility in stock market prediction research, where the choice of company count is strategically aligned with the study duration, data size, model complexity, data fusion strategies, and the intended level of analysis.

The stock market periods used in the reviewed studies exhibit a wide range in both duration and historical depth, reflecting the varied objectives and temporal resolutions of the predictive models. Some studies focused on short-term prediction windows lasting only a few months, such as between May 2018 and August 2018 or between June 2016 and March 2017, often aiming to capture high-frequency patterns or event-specific reactions. In contrast, some studies adopted extended time horizons, exceeding a decade (e.g., 1992–2016 or 2010–2021), to capture long-term market patterns and evaluate monthly fluctuations. A considerable portion of the reviewed works relied on mid-range windows of three to six years, which provided a practical balance between data availability and model complexity. Other studies concentrated on shorter and more recent periods, such as 2020–2022, to capture contemporary market behavior, particularly relevant to post-pandemic volatility. This variation in temporal scope demonstrates the adaptability of research designs and the importance of aligning data periods with model objectives, whether to examine macroeconomic cycles, market irregularities, or short-term investor sentiment.

Fig. 6 illustrates the distribution of data source combinations adopted across the reviewed studies. The majority of studies — 31 out of 44 — employed exactly two data sources, most commonly pairing historical stock prices with a single qualitative input such as financial news or social media sentiment. Ten studies incorporated three sources, while only three studies extended integration to four or more modalities. This pattern indicates that despite the theoretical advantages of multi-source fusion, the practical complexity of aligning and integrating heterogeneous data at scale continues to constrain the breadth of modality combinations explored in the literature.

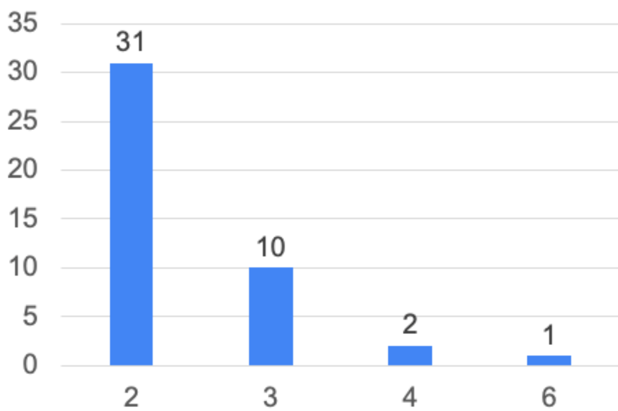

Fig. 6. Count of data sources used in each study.

In conclusion, recent literature on stock market prediction demonstrates an increasing trend toward multimodal integration of data sources, combining quantitative inputs (e.g., historical prices, macroeconomic indicators) with qualitative signals (e.g., news sentiment, social media activity). These contributions span multiple markets and regions, highlighting both the global relevance of the field and the necessity of multilingual processing, including English, Chinese, and Farsi. The unit of analysis ranges from single indices to hundreds of listed firms, while the temporal horizons extend from narrow short-term windows to multi-decade datasets. This diversity underscores the shift toward more comprehensive, context-aware, and flexible predictive frameworks.

TABLE II. SUMMARY OF REVIEWED STUDIES

Study	Country	Data Source	Duration (Year)	Number of Tickers
[20]	India	Reports	19	40
[21]	US	News	12	1
[22]	Indonesia	Reports, News, Google Trends	1	10
[23]	US, Chinese	Social Media	1.5	180
[24]	Chinese	Baidu Trends	11.25	1
[25]	Iran	Social Media	0.25	3
[26]	Australia	News, Social media	3.5	18
[27]	Taiwan	Google Trends	9.5	1
[28]	US, Pakistan	Social Media, Wikipedia	11	3
[29]	Ghana	Social Media, News, Google Trend, Forum	3	2
[30]	US, Pakistan	News, Social media	2.5	11
[31]	US, UK	Reports	6	na
[32]	Chinese	News	4	18
[33]	Chinese	News	3	37
[34]	US, Chinese	Forum	9	70
[35]	Pakistan	News	5	1
[36]	India	News, Social media	1	1
[37]	US	News	0.5	1
[38]	US	Social Media	12	10
[39]	US	News	1.15	12
[40]	US	News	5	6
[41]	US	News	7.5	1
[42]	Taiwan	Forum, Chip-Based Indicators	2.25	1
[43]	US	Reports, Social Media	5	33
[44]	US	News	10	3
[45]	Brazil	News, Social media	0.5	1
[46]	Chinese	Forum	2.5	30
[47]	US	Reports	5.25	240
[48]	Turkey	News, Social media	8	4
[49]	US	News	1	20
[50]	Poland	Forum	1	3
[51]	India	News	10	4
[52]	Taiwan	News, Institutional Investors	0.75	10
[53]	HongKong	News	5	12
[54]	US	Macroeconomic	24	13
[55]	Chinese	News	3	37
[56]	US	Social Media	2	88
[57]	Chinese	Social Media	2	1
[58]	Taiwan	News	6	30
[59]	US	News, Google Trends, Wikipedia	4	20
[60]	Japan	Social Media	8	1
[61]	Chinese	Reports Macroeconomic Indicators, Industry-Specific Metrics	13	3

B. RQ2: What Methodologies are used to Predict Stock Behavior?

Recent studies in AI-based stock market prediction consistently built on multimodal fusion frameworks that integrate quantitative and qualitative data within structured pipelines. These pipelines typically progress from data collection to feature engineering, model learning, and, in some cases, decision support. A commonly adopted design employs a two-branch or multi-branch architecture: one branch models market time series data, including prices, returns, and technical indicators, while the other uses the supporting data sources, such as extracting sentiment information from unstructured textual sources, such as financial news, social media, forums, or ESG narratives. The outputs from these branches are merged through a fusion layer for joint prediction. Fusion strategies vary, ranging from early fusion, where engineered features are concatenated before model training, to late fusion, which integrates the outputs of separately trained models, and hybrid fusion, in which specialized models feed into a final learner.

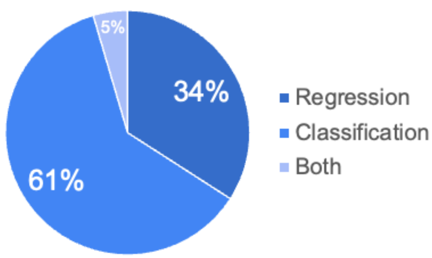

Fig. 7. Problem type.

Feature engineering remains as critical as model selection in these frameworks. For numerical market data, researchers compute extensive sets of technical indicators and integrate fundamental variables, accounting ratios, and macroeconomic metrics. Some studies leverage clustering techniques to group firms based on financial ratios, price patterns, or return profiles, thereby exploiting cross-sectional relationships. Textual feature extraction spans multiple approaches, including lexicon-based sentiment analysis (VADER, SentiWordNet, TextBlob, Azure), domain-specific dictionaries extracted from Wikipedia or ESG contexts, and transformer-based embeddings such as FinBERT. More advanced methods induce semantic and topical structures through TF-IDF with chi-square feature selection, latent Dirichlet allocation (LDA), joint sentiment–topic embeddings, contextual keyword augmentation using BERT, or semantic graph construction over named entities from curated Twitter accounts. To mitigate noise, several studies apply spam filtering, user credibility scoring, or restrict analysis to institutional investor sources. Temporal alignment is addressed by mapping late-day publications to the subsequent trading day. Finally, Google/Baidu Trends, forum thread counts, chip-based indicators (with genetic algorithm selection), and sector-specific factors (e.g., airlines) expand the exogenous signal set.

Learning models can be categorized into three main groups. First, classical statistical and machine learning algorithms (AR/ARIMA/SARIMAX, Linear/LASSO regression, SVM/SVR, decision trees, random forests, XGBoost, LightGBM, MKSVR, MVL-SVM) serve as benchmarks or components in larger systems. Second, deep models dominate temporal fusion tasks, with LSTM, Bi-LSTM, GRU, and Bi-GRU architectures frequently combined with CNNs for short-text classification or MLP/DNN structures for dense feature integration. Dual-branch architectures pairing FinBERT for text with LSTM for time series are particularly common, as are multi-step pipelines that integrate sentiment into TA/FI predictions. Third, reinforcement learning (RL) frameworks link prediction to trading execution, with multimodal features processed by sequence models feeding DQN-based policies for buy/hold decisions, explicitly embedding forecasting within action-oriented systems.

Prediction tasks reflect these methodological choices. As shown in Fig. 7, classification remains dominant, targeting directional movement (up/down, up/still/down), turning points, and volatility spikes, some works layering sentiment prediction first (bullish/bearish) and using it as an explanatory driver of direction, sometimes incorporating sentiment classification as an explanatory precursor. Regression tasks focus on predicting price levels, returns, or indices across various time horizons, from intraday intervals to quarterly forecasts. A subset of research advances beyond pure prediction to portfolio construction, applying top-K asset selection and benchmarking strategies against buy-and-hold baselines.

Evaluation protocols emphasize robustness. Studies typically benchmark against classical baselines (ARIMA/SARIMA, TI-only models, SVM, RF, MLP), conduct ablation testing (e.g., removing sentiment, removing TI, removing graph features) to quantify the incremental contribution of sentiment or technical/fundamental features, and several compare seed vs. augmented keyword sets, general vs customized lexicons, or pure-historical vs. full multimodal stacks. In trading-oriented settings, profitability and risk-adjusted returns serve as decision-centric measures.

Fig. 8 This taxonomy provides a focused merging of stock-prediction research, clarifying how modeling approaches, target variables, and evaluation metrics interact to shape methodological choices. It illustrates the field's shift toward deeper architectures and fusion strategies, while showing how different prediction goals—classification, regression, and volatility detection—demand distinct forms of analytical rigor. The structure emphasizes not only the diversity of techniques but also the logic connecting models to the outcomes they aim to forecast.

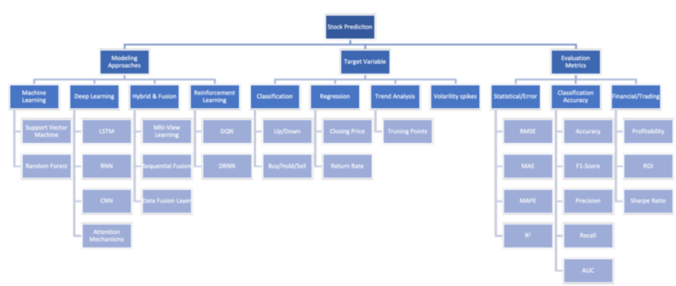

Fig. 8. Taxonomy of modeling approaches, target prediction types, and evaluation metrics used in stock market prediction research.

In summary, the methodological evolution of AI-based stock market prediction has shifted from isolated, single-source pipelines toward integrated, multimodal architectures that combine heterogeneous market signals with advanced sequence modeling and decision-aware design. The frontier of this research embeds forecasting within reinforcement learning frameworks, creating unified pipelines from data ingestion to trading action. Future directions include establishing standardized fusion benchmarks, adopting decision-relevant evaluation criteria, advancing multilingual NLP for diverse market contexts, and developing interpretable, reproducible systems capable of generalizing across markets and regimes.

C. RQ3: What are the different Performance Parameters used in Stock Market Prediction?

The selection of evaluation metrics in stock market prediction research is fundamentally shaped by whether the task is framed as a regression or classification problem. In regression-based studies, the objective is to forecast the precise numerical value of stock prices or indices—for example, predicting the next day's closing price—and performance is therefore assessed using error-focused measures such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). These performance metrics serve to quantify the gap between predicted and actual outcomes. RMSE accounts for both the magnitude and variance of prediction errors, while MAE offers a straightforward measure of the average deviation. MAPE is particularly useful in contexts where relative error is significant, as it enables comparisons across firms of different sizes and markets of varying scales. Collectively, these measures highlight the importance of minimizing forecast error, which is essential in financial time-series analysis where even small deviations may materially affect trading decisions. Fig. 8 shows the taxonomy of the evaluation metrics.

In classification-oriented research, the objective shifts toward predicting the directional movement of a stock—whether it rises, falls, or remains unchanged. Accuracy is the most frequently reported measure; however, its interpretive value diminishes under class imbalance, a common feature of financial datasets. To address this, studies often include Precision, Recall, and the F-score. Precision measures the proportion of correctly predicted upward movements, while Recall evaluates the ability to capture all true positives. The F-score balances both metrics harmoniously, ensuring that directional changes are identified with sensitivity and specificity. These indicators are critical in trading applications, where misclassifications can directly translate into financial loss.

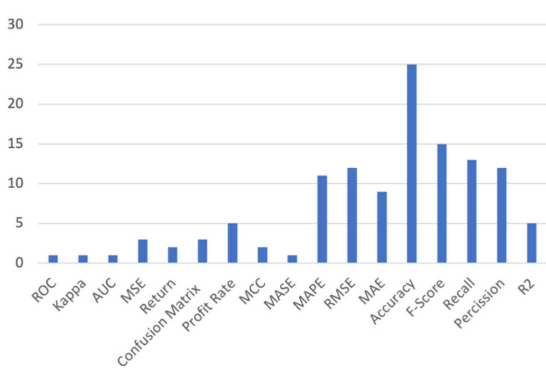

Fig. 9. Frequency of evaluation metrics employed in reviewed studies.

Beyond conventional statistical measures, some works employ advanced or hybrid evaluation frameworks that align predictive performance with investment outcomes. In regression, R^2 is occasionally used to assess explanatory power, while MSE complements RMSE by capturing error variance. In classification, ROC curves, AUC, Cohen's Kappa, MCC, and confusion matrices provide additional insight into model robustness under imbalanced scenarios. Furthermore, profit-oriented metrics such as total return [62], profit rate [32], [34], [45], and average gain [39] are utilized to directly link predictive accuracy with portfolio performance, bridging the gap between statistical evaluation and financial relevance.

In summary, the evaluation of stock market prediction models requires a multidimensional perspective that integrates statistical error minimization, directional accuracy, explanatory validity, and profitability. No single metric is sufficient in isolation; effective assessment must combine analytical rigor with decision-relevant measures to ensure that predictive models are both theoretically sound and practically applicable.

D. RQ4: What are the main difficulties in Predicting Stock Market behavior through Multi-Source Approaches?

It is important to recognize that while the integration of multiple data types can improve predictive performance, it simultaneously introduces substantial complexity. A recurring limitation identified in the literature is the fragmentation between data modalities. Many prior studies have focused exclusively on either qualitative or quantitative inputs, without establishing a unified analytical framework. This siloed orientation often results in biased or incomplete predictions. Moreover, combining heterogeneous sources—ranging from structured numerical data (e.g., historical stock prices, macroeconomic indicators) to unstructured text (e.g., financial news, forums, social media)—presents technical challenges. These include differences in data format, misaligned temporal resolutions, and inconsistent quality, all of which complicate feature selection, preprocessing, and data fusion. The unstructured nature of external textual data further exacerbates the difficulty, particularly when advanced context-aware NLP models are absent and keyword-based methods are employed without semantic depth.

Another significant challenge relates to methodological limitations and the continued reliance on single-source or single-modality approaches. Many studies emphasize technical indicators or historical price data alone, which risks overfitting and weak generalization. Similarly, models based solely on sentiment analysis often neglect temporal dynamics and market-specific dependencies. The lack of ensemble strategies and multimodal learning further undermines robustness. Even within deep learning applications, weaknesses persist in temporal modeling and the underutilization of reinforcement learning techniques, both of which are essential for capturing the dynamic and nonlinear structure of financial markets.

Linguistic and semantic issues also hinder the integration of multi-source data. Sentiment extraction across non-English languages—such as Chinese, Farsi, Portuguese, and Turkish—faces obstacles due to limited domain-specific lexicons, scarce NLP resources, and the semantic ambiguity inherent in financial discourse. For instance, aligning sentiment signals from Chinese or Persian texts with stock movements remains a major challenge due to linguistic complexity and cultural context. In addition, user-generated content on forums and social media often introduces noise, bias, and unreliability. Even when sentiment signals are correctly identified, they are rarely synchronized with real-time trading data, leading to a weak connection between extracted sentiment and actionable decision-making.

Finally, domain-specific and contextual constraints compound these challenges. Emerging markets, such as Brazil, Iran, and Pakistan, frequently experience limited data availability, low disclosure transparency, and a shortage of advanced modeling practices tailored to their structures. Moreover, external shocks—such as political instability or deviations from classical theories like the Efficient Market Hypothesis (EMH) and Random Walk Theory (RWT)—are difficult to incorporate into conventional models. Excessive irrelevant news, weak linkages between topics and sentiment, and the lack of evaluation frameworks aligned with investment strategies further limit practical value. Consequently, while multi-source integration offers richer insights, it requires advanced, context-aware, and semantically aligned models—capabilities that remain under development in current research.

V. Discussion

This review synthesizes findings from 44 primary studies published between 2018 and 2024 to derive comparative insights into heterogeneous-data strategies for stock market prediction. The evidence suggests that progress in this domain is shaped less by the selection of a single learning algorithm and more by the interaction of four interrelated elements: 1) the relevance of each data modality to the forecasting horizon, 2) the temporal alignment rules that determine what information is available at prediction time, 3) the fusion mechanism used to integrate heterogeneous signals, and 4) the evaluation design adopted to mitigate time-series leakage. When these elements are clearly specified and applied consistently, multimodal frameworks demonstrate credible predictive potential. Conversely, when alignment assumptions, fusion design, or validation procedures are insufficiently defined, reported performance gains become difficult to interpret and often exhibit limited transferability across markets, horizons, and datasets.

A central comparative outcome concerns the effectiveness of different modalities across forecasting horizons. Price-based variables and technical indicators remain the predominant inputs due to their dense availability, precise time-stamping, and natural compatibility with short-horizon objectives such as next-day movement or return prediction. Textual information—particularly financial news—emerges as the most common complementary source, as it provides event-driven context that may not be immediately reflected in historical prices. However, the contribution of sentiment derived from textual sources is not consistent across studies. Its predictive value is most evident when the target variable is sensitive to information shocks and when textual inputs are aligned with the market's assimilation window. If publication time, trading hours, and aggregation windows are not handled explicitly, textual features may unintentionally incorporate forward-looking information, producing inflated performance that does not reflect a genuine predictive signal. Accordingly, sentiment is best treated as a time-constrained modality whose utility depends on disciplined temporal alignment rather than as a universally beneficial feature. More broadly, the evidence indicates that the effectiveness of multimodal approaches depends less on increasing the number of inputs and more on ensuring that each modality is temporally consistent with the target definition and integrated in a leakage-safe manner.

The reviewed literature also indicates that most multimodal pipelines are constructed using two modalities—most commonly price data combined with textual information—whereas fewer studies incorporate three or more sources, such as fundamentals, search trends, or alternative external signals. This pattern reflects a practical trade-off between informational richness and integration complexity. As additional modalities are introduced, the dominant challenges shift from model expressiveness to data engineering considerations, including mismatched sampling frequencies, missing values, source noise, and the tendency for high-dimensional representations (notably text embeddings) to dominate learning. These observations suggest that multimodal stock prediction is fundamentally an alignment and integration problem, in which pipeline design and data handling often exert greater influence on downstream performance than incremental increases in model complexity.

Fusion strategy is a key determinant of robustness, yet it is frequently described at a high level without sufficient implementation detail to support rigorous cross-study comparison. Feature-level (early) fusion enables a unified learner to capture cross-modal interactions, but it is highly sensitive to temporal misalignment, scaling discrepancies, and leakage introduced during feature construction. Decision-level (late) fusion is generally more robust under heterogeneity, as it models each modality independently, reduces sensitivity to frequency mismatch, and supports modular handling of missing inputs. Hybrid fusion strategies—such as multi-branch architectures, gating mechanisms, and attention-based integration—provide a compromise by learning modality-specific representations while enabling interaction at a learned fusion stage. Although such designs are well aligned with market dynamics, where source relevance may vary across regimes, they require stronger empirical support through ablation testing to ensure that reported gains arise from multimodal integration rather than increased parameterization.

Model selection follows the same comparative logic. Classical machine-learning approaches remain competitive when datasets are limited or indicators are well engineered, often providing stable and efficient baselines. Deep sequence models are widely used when temporal dependencies are central; however, their advantages depend on adequate data volume and leakage-safe validation. For textual modeling, architectural selection should reflect the characteristics of financial language and the nature of the text source. In particular, bi-directional recurrent architectures are commonly motivated in sentiment settings where meaning depends on surrounding context (e.g., negation and contrast), whereas convolutional encoders are effective for short headline-style inputs where local patterns dominate and computational efficiency is desirable. Transformer-based models are most appropriate when richer contextual representations are required or when domain-specific semantics influence interpretation, and their practical benefit is typically strongest when textual inputs are sufficiently informative and consistently aligned with market movement. Overall, the evidence supports selecting architectures based on data–task alignment rather than assuming that deeper models are inherently superior.

Evaluation practices account for a substantial portion of the variation in reported performance. Given the time-series nature of stock prediction, experimental validity depends critically on preserving chronological integrity. Random splits, overlapping feature windows, or ambiguous text-to-label alignment can introduce forward-looking information into training and produce overly optimistic estimates. This risk is amplified in multimodal designs, where textual data creates additional opportunities for subtle leakage through timestamp handling, particularly when after-hours news is not handled explicitly or when text aggregation windows overlap with the target horizon. Comparability is further reduced by heterogeneous target definitions, inconsistent metric selection, and incomplete reporting of baselines and validation procedures. Therefore, time-aware validation, explicit modality timing rules, and ablation against strong unimodal baselines should be regarded as essential practices for substantiating claims of multimodal advantage.

Finally, generalization across markets and languages remains a practical challenge. While many studies focus on English-language sources and information-rich markets, multilingual and emerging-market contexts differ in disclosure practices, liquidity characteristics, and information ecosystems. In such settings, sentiment quality and feature stability vary substantially, and models trained in one market or language may not transfer without adaptation. This reinforces the need for transparent reporting of data provenance, preprocessing steps, and alignment assumptions, as these factors determine whether observed gains reflect market-specific artifacts or broadly applicable methodological improvements.

In summary, the evidence from the reviewed studies indicates that heterogeneous-data stock prediction is promising but methodologically sensitive. Reliable gains are most consistently observed when modality selection is aligned with the prediction horizon, fusion strategy is matched to data heterogeneity, and evaluation protocols explicitly mitigate time-series leakage. Future progress is therefore expected to depend less on increasing model complexity and more on improved methodological rigor, including clear definitions of information availability at prediction time, implementation-level specification of fusion mechanisms, leakage-safe validation, and systematic ablation to quantify the incremental value of each modality.

VI. Conclusion

This systematic review provides a consolidated understanding of how heterogeneous data, machine-learning methodologies, and evaluation practices have evolved in stock-market prediction research. The evidence shows that models leveraging multimodal information—technical, fundamental, and sentiment-based—tend to achieve superior predictive performance compared with single-source approaches. However, the literature is marked by methodological variability and an absence of unified evaluation standards, which limits comparability and weakens empirical generalization.

Overall, this review highlights both the potential and the challenges of multimodal stock prediction. It contributes by organizing current knowledge, identifying methodological patterns, and clarifying research gaps that future work must address. To advance the field, upcoming studies should adopt more transparent protocols, incorporate robust validation strategies, and evaluate models using metrics that reflect real financial impact rather than relying solely on statistical accuracy indicators.

VII. Future Work

Several directions merit attention in future research. Standardized benchmarking frameworks should be developed to enable fair cross-study comparison of fusion approaches, addressing the inconsistent evaluation practices identified in this review. Additionally, the development of domain-specific NLP resources for non-English financial markets and the incorporation of model interpretability as a first-class evaluation criterion represent promising avenues for advancing the field, given the growing emphasis on transparency in AI-driven financial decision-making.

References

[1] S. Islam, Md. S. Sikder, Md. F. Hossain, and P. Chakraborty, "Predicting the daily closing price of selected shares on the Dhaka Stock Exchange using machine learning techniques," SN Business & Economics, vol. 1, no. 4, pp. 1–16, 2021, doi: 10.1007/s43546-021-00065-6.

[2] A. Saini and A. Sharma, "Predicting the Unpredictable: An Application of Machine Learning Algorithms in Indian Stock Market," Annals of Data Science, vol. 9, no. 4, pp. 791–799, 2022, doi: 10.1007/s40745-019-00230-7.

[3] A. Pathak and N. P. Shetty, Indian Stock Market Prediction Using Machine Learning and Sentiment Analysis, vol. 711. Springer Singapore, 2019. doi: 10.1007/978-981-10-8055-5_53.

[4] E. F. Fama, "Efficient Market Hypothesis: A Review of Theory and Empirical Work," J. Finance, vol. 25, no. 2, pp. 383–417, 1970.

[5] D. Shah, H. Isah, and F. Zulkernine, "Stock market analysis: A review and taxonomy of prediction techniques," International Journal of Financial Studies, vol. 7, no. 2, 2019, doi: 10.3390/ijfs7020026.

[6] L. N. Mintarya, J. N. M. Halim, C. Angie, S. Achmad, and A. Kurniawan, "Machine learning approaches in stock market prediction: A systematic literature review," in Procedia Computer Science, Elsevier B.V., 2022, pp. 96–102. doi: 10.1016/j.procs.2022.12.115.

[7] D. Kumar, P. K. Sarangi, and R. Verma, "A systematic review of stock market prediction using machine learning and statistical techniques," in Materials Today: Proceedings, Elsevier Ltd, 2020, pp. 3187–3191. doi: 10.1016/j.matpr.2020.11.399.

[8] N. Lumoring, D. Chandra, and A. A. S. Gunawan, "A Systematic Literature Review: Forecasting Stock Price Using Machine Learning Approach," in 2023 International Conference on Data Science and Its Applications, ICoDSA 2023, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 129–133. doi: 10.1109/ICoDSA58501.2023.10277318.

[9] S. Hingane, M. Bhate, H. Patil, and P. Ukhalkar, "Literature Review on Stock Price Predictions with reference to Machine Learning Approaches and Statistical Methods," 2025. [Online]. Available: https://internationalpubls.com.

[10] R. Jain and R. Vanzara, “Emerging Trends in AI-Based Stock Market Prediction: A Comprehensive and Systematic Review †,” Engineering Proceedings, vol. 56, no. 1, 2023, doi: 10.3390/ASEC2023-15965.

[11] D. Shah, H. Isah, and F. Zulkernine, “Stock market analysis: A review and taxonomy of prediction techniques,” 2019, MDPI Multidisciplinary Digital Publishing Institute. doi: 10.3390/ijfs7020026.

[12] N. Rouf et al., “Stock market prediction using machine learning techniques: A decade survey on methodologies, recent developments, and future directions,” Nov. 01, 2021, MDPI. doi: 10.3390/electronics10212717.

[13] O. Bustos and A. Pomares-Quimbaya, “Stock market movement forecast: A Systematic review,” Oct. 15, 2020, Elsevier Ltd. doi: 10.1016/j.eswa.2020.113464.

[14] S. Usmani and J. A. Shamsi, “News sensitive stock market prediction: Literature review and suggestions,” PeerJ Comput. Sci., vol. 7, pp. 1–36, 2021, doi: 10.7717/PEERJ-CS.490.

[15] A. Ehsan, S. Habib, and A. Sohail, “Financial News Sentiment Analysis Using NLP and Machine Learning for Asset Price Prediction: A Systematic Review,” VFAST Transactions on Software Engineering, vol. 13, no. 3, pp. 279–308, Sep. 2025, doi: 10.21015/vtse.v13i3.2165.

[16] R. W. A. Dipura, F. I. Maulana, and Yulianto, “A Systematic Literature Review For Stock Price Prediction with Machine Learning,” in Proceedings of the 2025 11th International Conference on Communication and Signal Processing, ICCSP 2025, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 1263–1268. doi: 10.1109/ICCSP64183.2025.11088479.

[17] I. K. Nti, A. F. Adekoya, and B. A. Weyori, “A systematic review of fundamental and technical analysis of stock market predictions,” Artif. Intell. Rev., vol. 53, no. 4, pp. 3007–3057, 2020, doi: 10.1007/s10462-019-09754-z.

[18] B. A. Kitchenham and S. Charters, Guidelines for performing Systematic Literature Reviews in Software Engineering, no. January. 2007.

[19] M. Alkhamees, S. Alsaleem, M. Al-Qurishi, M. Al-Rubaian, and A. Hussain, “User trustworthiness in online social networks: A systematic review,” Appl. Soft Comput., vol. 103, p. 107159, 2021, doi: 10.1016/j.asoc.2021.107159.

[20] J. Kaur and K. Dharni, “Data mining–based stock price prediction using hybridization of technical and fundamental analysis,” Data Technologies and Applications, vol. 57, no. 5, pp. 780–800, Nov. 2023, doi: 10.1108/DTA-04-2022-0142.

[21] K. Puh and M. Bagić Babac, “Predicting stock market using natural language processing,” American Journal of Business, vol. 38, no. 2, pp. 41–61, May 2023, doi: 10.1108/ajb-08-2022-0124.

[22] S. Agusta, F. Rakhman, J. H. Mustakini, and S. Wijayana, “Enhancing the accuracy of stock return movement prediction in Indonesia through recent fundamental value incorporation in multilayer perceptron,” Asian Journal of Accounting Research, vol. 9, no. 4, pp. 358–377, Aug. 2024, doi: 10.1108/AJAR-01-2024-0006.

[23] Y. Shi, Y. Zheng, K. Guo, and X. Ren, “Stock movement prediction with sentiment analysis based on deep learning networks,” Concurr. Comput., vol. 33, no. 6, Mar. 2021, doi: 10.1002/cpe.6076.

[24] X. Tang, N. Lei, M. Dong, and D. Ma, “Stock Price Prediction Based on Natural Language Processing1,” Complexity, vol. 2022, 2022, doi: 10.1155/2022/9031900.

[25] A. Hatefi Ghahfarrokhi and M. Shamsfard, “Tehran stock exchange prediction using sentiment analysis of online textual opinions,” Intelligent Systems in Accounting, Finance and Management, vol. 27, no. 1, pp. 22–37, Jan. 2020, doi: 10.1002/isaf.1465.

[26] B. J. Vanstone, A. Gepp, and G. Harris, “Do news and sentiment play a role in stock price prediction?,” Applied Intelligence, vol. 49, no. 11, pp. 3815–3820, 2019, doi: 10.1007/s10489-019-01458-9.

[27] M. H. Fan, M. Y. Chen, and E. C. Liao, “A deep learning approach for financial market prediction: utilization of Google trends and keywords,” Granular Computing, vol. 6, no. 1, pp. 207–216, 2021, doi: 10.1007/s41066-019-00181-7.

[28] W. Khan, U. Malik, M. A. Ghazanfar, M. A. Azam, K. H. Alyoubi, and A. S. Alfakeeh, “Predicting stock market trends using machine learning algorithms via public sentiment and political situation analysis,” Soft comput., vol. 24, no. 15, pp. 11019–11043, 2020, doi: 10.1007/s00500-019-04347-y.

[29] I. K. Nti, A. F. Adekoya, and B. A. Weyori, “A novel multi-source information-fusion predictive framework based on deep neural networks for accuracy enhancement in stock market prediction,” J. Big Data, vol. 8, no. 1, 2021, doi: 10.1186/s40537-020-00400-y.

[30] W. Khan, M. A. Ghazanfar, M. A. Azam, A. Karami, K. H. Alyoubi, and A. S. Alfakeeh, “Stock market prediction using machine learning classifiers and social media, news,” J. Ambient Intell. Humaniz. Comput., vol. 13, no. 7, pp. 3433–3456, 2022, doi: 10.1007/s12652-020-01839-w.

[31] T.-Y. Hsu, “Machine learning applied to stock index performance enhancement,” Journal of Banking and Financial Technology, vol. 5, no. 1, pp. 21–33, Jun. 2021, doi: 10.1007/s42786-021-00025-6.

[32] X. Chen, K. Hirota, Y. Dai, and Z. Jia, “A model fusion method based on multi-source heterogeneous data for stock trading signal prediction,” Jul. 13, 2022. doi: 10.21203/rs.3.rs-1576411/v1.

[33] W. Long, J. Gao, K. Bai, and Z. Lu, “A hybrid model for stock price prediction based on multi-view heterogeneous data,” Financial Innovation, vol. 10, no. 1, Dec. 2024, doi: 10.1186/s40854-023-00519-w.

[34] W. Liu, Y. Gu, and Y. Ge, “Multi-factor stock trading strategy based on DQN with multi-BiGRU and multi-head ProbSparse self-attention,” Applied Intelligence, vol. 54, no. 7, pp. 5417–5440, Apr. 2024, doi: 10.1007/s10489-024-05463-5.

[35] A. Zaffar and S. M. A. Hussain, “Modeling and prediction of KSE – 100 index closing based on news sentiments: an applications of machine learning model and ARMA (p, q) model,” Multimed. Tools Appl., vol. 81, no. 23, pp. 33311–33333, 2022, doi: 10.1007/s11042-022-13052-2.

[36] N. Das, B. Sadhukhan, T. Chatterjee, and S. Chakrabarti, “Effect of public sentiment on stock market movement prediction during the COVID-19 outbreak,” Soc. Netw. Anal. Min., vol. 12, no. 1, Dec. 2022, doi: 10.1007/s13278-022-00919-3.

[37] Y. Li and Y. Pan, “A Novel Ensemble Deep Learning Model for Stock Prediction Based on Stock Prices and News,” Jul. 2020, [Online]. Available: http://arxiv.org/abs/2007.12620

[38] A. Bodaghi and J. J. H. Zhu, “Using dynamic semantic structure of news flow to enhance financial forecasting: a twelve-year study on twitter news channels,” Multimed. Tools Appl., vol. 84, no. 24, pp. 28191–28223, Jul. 2025, doi: 10.1007/s11042-024-20274-z.

[39] P. Chen, Z. Boukouvalas, and R. Corizzo, “A deep fusion model for stock market prediction with news headlines and time series data,” Neural Comput. Appl., vol. 36, no. 34, pp. 21229–21271, Dec. 2024, doi: 10.1007/s00521-024-10303-1.

[40] Z. Wang, Z. Hu, · Fang Li, and S.-B. Ho, “Learning-Based Stock Market Trending Analysis by Incorporating Social Media Sentiment Analysis,” 2023. Accessed: Aug. 17, 2025. [Online]. Available: https://www.researchgate.net/publication/351177658_Learning-Based_Stock_Market_Trending_Analysis_by_Incorporating_Social_Media_Sentiment_Analysis

[41] H. Lee, J. H. Kim, and H. S. Jung, “Deep-learning-based stock market prediction incorporating ESG sentiment and technical indicators,” Sci. Rep., vol. 14, no. 1, May 2024, doi: 10.1038/s41598-024-61106-2.

[42] J. Y. Huang, C. L. Tung, and W. Z. Lin, “Using Social Network Sentiment Analysis and Genetic Algorithm to Improve the Stock Prediction Accuracy of the Deep Learning-Based Approach,” International Journal of Computational Intelligence Systems, vol. 16, no. 1, Dec. 2023, doi: 10.1007/s44196-023-00276-9.

[43] M. A. Arauco Ballesteros and E. A. Martínez Miranda, “Stock Market Forecasting Using a Neural Network Through Fundamental Indicators, Technical Indicators and Market Sentiment Analysis,” Comput. Econ., 2024, doi: 10.1007/s10614-024-10711-4.

[44] J. Readshaw and S. Giani, “Using company-specific headlines and convolutional neural networks to predict stock fluctuations,” Neural Comput. Appl., vol. 33, no. 24, pp. 17353–17367, 2021, doi: 10.1007/s00521-021-06324-9.

[45] A. E. de Oliveira Carosia, G. P. Coelho, and A. E. A. da Silva, “Investment strategies applied to the Brazilian stock market: A methodology based on Sentiment Analysis with deep learning,” Expert

Syst. Appl., vol. 184, no. June, p. 115470, 2021, doi: 10.1016/j.eswa.2021.115470.

[46] N. Jing, Z. Wu, and H. Wang, "A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction," Expert Syst. Appl., vol. 178, no. March, p. 115019, 2021, doi: 10.1016/j.eswa.2021.115019.

[47] J. Vásquez Sáenz, F. M. Quiroga, and A. F. Bariviera, "Data vs. information: Using clustering techniques to enhance stock returns forecasting," International Review of Financial Analysis, vol. 88, Jul. 2023, doi: 10.1016/j.irfa.2023.102657.

[48] Y. E. Akdogan and A. Anbar, "More than just sentiment: Using social, cognitive, and behavioral information of social media to predict stock markets with artificial intelligence and big data," Borsa Istanbul Review, vol. 24, pp. 61–82, Dec. 2024, doi: 10.1016/j.bir.2024.12.003.

[49] A. Picasso, S. Merello, Y. Ma, L. Oneto, and E. Cambria, "Technical analysis and sentiment embeddings for market trend prediction," Expert Syst. Appl., vol. 135, pp. 60–70, 2019, doi: 10.1016/j.eswa.2019.06.014.

[50] G. Wojarnik, "Sentiment analysis as a factor included in the forecasts of price changes in the stock exchange," Procedia Comput. Sci., vol. 192, pp. 3176–3183, 2021, doi: 10.1016/j.procs.2021.09.090.

[51] J. Maqbool, P. Aggarwal, R. Kaur, A. Mittal, and I. A. Ganaie, "Stock Prediction by Integrating Sentiment Scores of Financial News and MLP-Regressor: A Machine Learning Approach," in Procedia Computer Science, Elsevier B.V., 2022, pp. 1067–1078. doi: 10.1016/j.procs.2023.01.086.

[52] H. C. Wang, W. C. Hsiao, and R. S. Liou, "Integrating technical indicators, chip factors and stock news for enhanced stock price predictions: A multi-kernel approach," Asia Pacific Management Review, vol. 29, no. 3, pp. 292–305, Sep. 2024, doi: 10.1016/j.apmrv.2023.10.001.

[53] X. Li, P. Wu, and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong," Inf. Process. Manag., vol. 57, no. 5, p. 102212, 2020, doi: 10.1016/j.ipm.2020.102212.

[54] B. Weng et al., "Macroeconomic Indicators Alone can Predict the Monthly Closing Price of Major U.S. Indices: Insights from Artificial Intelligence, Time-Series Analysis and Hybrid Models," 2018.

[55] W. Long, J. Gao, and M. Guo, "Comparative research on multi-source heterogeneous data fusion technologies," in Procedia Computer Science, Elsevier B.V., 2024, pp. 1089–1095. doi: 10.1016/j.procs.2024.08.198.

[56] H. Xu, L. Chai, Z. Luo, and S. Li, "Stock movement predictive network via incorporative attention mechanisms based on tweet and historical prices," Neurocomputing, vol. 418, pp. 326–339, 2020, doi: 10.1016/j.neucom.2020.07.108.

[57] W. Chen, C. K. Yeo, C. T. Lau, and B. S. Lee, "Leveraging social media news to predict stock index movement using RNN-boost," Data Knowl. Eng., vol. 118, pp. 14–24, 2018, doi: 10.1016/j.datak.2018.08.003.

[58] P. Y. Hao, C. F. Kung, C. Y. Chang, and J. B. Ou, "Predicting stock price trends based on financial news articles and using a novel twin support vector machine with fuzzy hyperplane," Appl. Soft Comput., vol. 98, Jan. 2021, doi: 10.1016/j.asoc.2020.106806.

[59] B. Weng, L. Lu, X. Wang, F. M. Megahed, and W. Martinez, "Predicting short-term stock prices using ensemble methods and online data sources," Expert Syst. Appl., vol. 112, pp. 258–273, 2018, doi: 10.1016/j.eswa.2018.06.016.

[60] K. Ueda et al., "SSCDV: Social media document embedding with sentiment and topics for financial market forecasting," Expert Syst. Appl., vol. 245, Jul. 2024, doi: 10.1016/j.eswa.2023.122988.

[61] X. Xu, Y. Zhang, C. A. McGrory, J. Wu, and Y. G. Wang, "Forecasting stock closing prices with an application to airline company data," Data Science and Management, vol. 6, no. 4, pp. 239–246, Dec. 2023, doi: 10.1016/j.dsm.2023.09.005.

[62] Y. F. Chen and S. H. Huang, "Sentiment-influenced trading system based on multimodal deep reinforcement learning," Appl. Soft Comput., vol. 112, p. 107788, 2021, doi: 10.1016/j.asoc.2021.107788.