# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Introduction

Over the years, traders, economists, and policy-makers have all dreamed of nailing down tomorrow's stock price. Dozens of classic models once dominated that chase, from shifting averages to dividend discounts. extra currently, breakthroughs in machine mastering and natural language processing have opened the door to messier, tougher-to-quantify information. Tweets, earnings call transcripts, and blog posts can now feed directly into algorithms. This chapter will discuss about research landscape, sentiment, and deep-learning techniques. It lays out the tools others have used, points out what still feels missing, and explains why this study is conducted.

## 2.2 Stock Market Prediction: An Overview

Predicting stock market behaviour requires analysts to look forward into tomorrow's prices by weighing whatever historical price trends, financial benchmarks, and external pressures they can muster. for many years, the ones forecasts were built commonly on fairly truthful statistical workouts like Autoregressive integrated moving average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH). both fashions hinge on the assumption that the underlying records keeps a few degree of linearity and stationarity. Although they retain clear advantages in interpretability and computational speed, they struggle when faced with the nonlinear spikes, regime shifts, and multidimensional clutter that financial markets routinely generate (Alamu & Siam, 2024).

The scene has shifted dramatically during the last ten years as machine learning and, more recently, deep learning have claimed centre stage in financial forecasting. traditional algorithms—help Vector Machines, choice timber, and ensemble techniques consisting of Gradient Boosting—have time and again outperformed older techniques by way of uncovering diffused systems buried in tangled facts. aid Vector Machines, as an instance, are in particular popular because they restriction overfitting even as hopefully navigating curved decision limitations (Geetha et al., 2024). Crucially, those techniques can accommodate an eclectic blend of indicators: the whole thing from rate momentum and moving averages to macroeconomic releases and social media sentiment can be folded into a single version within the desire of lifting prediction precision.

Recurrent Neural Networks (RNNs) and their more advanced offspring, especially Long-Short Term Memory (LSTM) networks, have swiftly installed themselves as powerful gear for modeling sequential economic data. The LSTM architecture excels in time-collection forecasting mainly due to the fact it is able to successfully remember information over prolonged intervals whilst largely fending off the vanishing gradient problem that hampers traditional RNNs (Ansah et al., 2022). This advantage has been empirically established; in one current head-to-head assessment, LSTM outstripped both ARIMA and help Vector gadget (SVM) fashions at the imply Squared errors (MSE) metric throughout several markets, ranging from equities to cryptocurrencies such as Bitcoin (Al-Alawi & Alshakhoori, 2024).

Alongside these stand-alone networks, hybrid frameworks that meld different machine-learning techniques are increasingly popular in the field. By fusing LSTM with techniques like LightGBM or by employing CNN-BiLSTM configurations, researchers can effectively extract both spatial and temporal patterns from the data. Such integrative methods frequently deliver superior forecasting accuracy when pitted against any single architecture (Yu, 2024). Moreover, Transformer-primarily based designs, such as the Temporal Fusion Transformer (TFT) and the LSTM-mTrans-MLP, have these days demonstrated marked gains in prediction accuracy, particularly when tasked with high-frequency datasets like the S&P 500 index (Kabir et al., 2025).

Machine-learning (ML) and deep-learning (DL) systems have come a long way, yet they still grapple with several practical drawbacks. For one, the hazard of overfitting looms huge, particularly in rapid-moving monetary environments wherein past styles quick lose relevance. further, these techniques often demand sizable education datasets and hefty computing electricity. at the same time as deep architectures can outperform simpler procedures, their opaque selection-making—frequently known as the “black box” hassle—makes it hard for buyers and analysts to comprehend why a given name changed into made (Supendi et al., 2024).

In contrast, more traditional forecasting methods, such as autoregressive integrated moving average (ARIMA) and exponential smoothing, continue to be useful when the need for clarity and computational speed trumps raw power. These time-tested models shine in stable, short-horizon projections, allowing users to follow their reasoning step by step. Yet as turbulence in the markets rises, they often fall behind, ceding ground to ML techniques designed to exploit noisy, non-stationary data.

The latest studies stress the value of melding disparate data streams—real-time news feeds, social-media sentiment, option-implied volatility—into a single predictive framework. Evidence gathered across several investigations shows that when such sentiment-rich information feeds through long short-term memory (LSTM) networks or convolutional neural networks (CNNs), forecasting accuracy improves noticeably by accounting for investor mood and attitude as drivers of price movement (Kapgate & Chaturvedi, 2025).

To summarise, traditional statistical techniques remain the bedrock of financial forecasting, yet machine-learning and deep-learning methods now provide flexible alternatives that can better accommodate the intricate dynamics of contemporary markets. The rise of hybrid architectures and transformer-based systems points to a promising trajectory for improving accuracy and informing decision-making in equity analysis.

### 2.3 Role of Sentiment in Financial Market

Investor sentiment, appearing as a powerful mental undercurrent, frequently steers costs in guidelines that fundamental variables can't absolutely justify. within the framework of behavioural finance, sentiment is perceived as a collective emotional kingdom—moulded with the aid of headlines, professional remark, and social media chatter—that frequently triggers overreactions, distortions, and spikes in volatility (Nyakurukwa & Seetharam, 2024). In light of this insight, students have begun to embed sentiment rankings extracted from text into forecasting models, hoping to refine predictive electricity and render marketplace behaviour more understandable.

Sentiment analysis has become an integral part of contemporary financial research, using natural language processing to extract and measure the views embedded in market-related texts. while early efforts regularly trusted primary sentiment lexicons, the field has moved in the direction of greater effective architectures consisting of Bidirectional Encoder Representations from Transformers (BERT), XLNet, and diverse hybrid deep-studying frameworks. A latest contribution by Fu and Zhang (2024) illustrates this fashion; the authors developed a BERT-derived model that converts streams of financial documents into sentiment time series, which are then paired with traditional technical indicators and fed into a two-layer long short-term memory (LSTM) network for stock forecasting. Their results revealed a clear edge over pure quantitative models, underscoring how discernible emotional signals can enhance predictive accuracy.

The blessings of sentiment evaluation enlarge beyond mere type of textual content polarity. with the aid of incorporating dictionaries crafted specifically for the finance region, researchers have all started to provide finer-grained insights. Wang (2023) leveraged SnowNLP alongside a bespoke sentiment lexicon designed for chinese inventory forums, in the end creating an index that, when mixed with past charge actions, appreciably sharpened trend forecasts. In a related study, Agarwal and Gupta (2024) showed that large language models like Llama 2, once fine-tuned on financial datasets, reached over 89 percent accuracy in unraveling sentiment subtleties, thereby leaving traditional lexicon-driven approaches at a disadvantage.

Sentiment data is increasingly being harnessed outside the traditional confines of financial reporting. Social platforms, and in particular Twitter and Reddit, now serve as immediately barometers for investor sentiment. Osman (2023) proven that when Twitter mood is fused with historical price facts thru an ensemble approach, subsequent-hour inventory forecasts attain an accuracy of seventy four.3 in step with cent. Their research observed that social sentiment, rather than structured news, exerts a greater influence on U.S. intraday price swings. Conversely, Nyakurukwa and Seetharam (2024) found that in emerging and frontier markets, signals from official news channels carry more weight than social chatter, highlighting a geographic discrepancy in sentiment's power.

As analytical models evolve, so too does their capacity for prediction. Jiang et al. (2024) unveiled AEformer, a transformer-based architecture that leverages asymmetric embeddings and tailored attention layers to track rapid sentiment shifts typical of high-frequency trading. In parallel, Li and Hu (2024) crafted a hybrid framework which extracts sentiment with XLNet, layers in popularity data and technical indicators, and channels the output into a dual BiLSTM-highway pipeline. This setup reported tangible gains not only in short-term accuracy but also in risk-adjusted performance metrics.

Even with recent progress in the field, sentiment analysis still faces several obstacles. Models performing these analyses often struggle when confronted with poor-quality data, sarcastic remarks, ambiguous phrasing, or language that is specific to a particular industry. The rapid turnaround of events in financial markets makes it even harder to interpret sentiment information in real time. On top of that, measuring how sentiment actually moves prices is tricky in markets that are already very efficient, since arbitrage traders quickly incorporate any public information, they see (Kaveri et al., 2025).

## 2.4 Sentiment Analysis in Stock Prediction

### 2.4.1 Sentiment Analysis Techniques

At its core, sentiment analysis is an essential challenge inside natural language processing (NLP) that targets to spot, pull aside, and label the opinionated pieces of text. while implemented to stock-marketplace forecasting, it seeks to turn investor moods and mental cues located in news articles, profits releases, or social-media chatter into numbers that models can work with. Those numbers are then fed into predictive systems as extra features, hopefully leading to better forecasts. To achieve this, practitioners typically lean on three broad method families: lexicon-based rules, traditional machine-learning classifiers, and modern deep-learning networks.

Lexicon-driven sentiment analysis depends on curated lists of words each assigned a positive or negative score. Because these tools are straightforward to build and light on processing power, they remain popular. Yet they frequently misread industry jargon present in financial writing. Broad lexicons such as VADER and AFINN routinely misclassify terms like “bullish” or “overweight,” failing to acknowledge that their meanings shift sharply in a trading context (Nagendra et al., 2024). To address this weakness, researchers create dedicated financial dictionaries. The catch is that assembling and updating such lists demands considerable manual effort, a chore that hampers growth. In response, recent studies have unveiled eXplainable Lexicons, or XLex, which marry transformer-driven contextual embeddings to traditional rule-based glossaries. The goal is to fuse high accuracy with clear reasoning (Rizinski et al., 2024).

Machine-learning pipelines, by contrast, build classifiers on datasets marked with sentiment labels. Popular algorithms—Naïve Bayes, Logistic Regression, Random Forest, Support Vector Machine—are trained after text is converted into numerical form via TF-IDF or Bag-of-Words. Umar (2023) reported that both Naïve Bayes and Logistic Regression gauge sentiment in financial news with impressive precision when features are extracted at the document level. Still, these methods can be hungry for manual feature tuning and vast annotated collections, making them costly in both time and data.

In recent years, deep learning approaches have transformed sentiment analysis by removing the need for manual feature engineering and by exposing subtle meaning that earlier methods often missed. amongst the earlier successes, recurrent neural networks—especially the long short-time period reminiscence units and their slimmer counterpart, the Gated Recurrent Unit—have proven powerful at following the chain of phrases over the years, which makes them a natural fit for tasks in financial sentiment analysis (Paulraj, 2024). Researchers soon began stacking layers and adding convolutional filters, so the bidirectional LSTM-CNN combination could simultaneously pick up short-range phrases and broader sentence structure, producing noticeable gains in classification accuracy (Aluvala et al., 2023).

Lately, transformer architectures have taken the lead, and models such as BERT (Bidirectional Encoder Representations from Transformers) routinely set the bar higher for many benchmarking datasets. For instance, Lin and Wang (2024) paired BERT with a lightweight transformer layer that brings together polarity scores extracted from news headlines and features extracted from transaction logs. Their hybrid system reached an impressive 60% forecasting accuracy for index movements. Despite these breakthroughs, the heavy appetite for processing power and data remains a concern that researchers must manage when scaling such models for real-world applications.

Hybrid modelling approaches are attracting considerable attention as researchers seek to leverage the complementary advantages of distinct analytical techniques. One promising example is the recently proposed MS-SSA-LSTM architecture, which fuses standards from swarm intelligence, a curated sentiment index, and a multi-layer deep learning framework to deliver brief-time period forecasts for inventory charges. preliminary opinions propose that this composite system consistently outperforms traditional single-algorithm baselines (Madhuri et al., 2024). Another noteworthy development pairs knowledge graphs with attention-enhanced long short-term memory (LSTM) networks, thereby refining the model's ability to encode financial news and its intricate ties to subsequent market dynamics (Zhang, 2023).

Yet, progress in these areas is far from straightforward. Financial commentary—whether published articles, press releases, or tweets— frequently employs sarcasm, area of interest terminology, and complex syntactic bureaucracy that prevent dependable sentiment extraction. moreover, even when preprocessing pipelines succeed in cleaning the textual content, the signals themselves remain volatile and, at times, mutually exclusive, especially when the input stream aggregates markedly different platforms such as formal news sites and informal social feed.

### 2.4.2 Application in Financial News

Analysing the emotional tone of financial news headlines has increasingly proven itself useful for anticipating stock price movements, especially in emerging markets like Malaysia, where prices can react with particular intensity to newly released information. By mining the sentiment woven into news stories, analysts supplement standard quantitative inputs—such as past prices and trading volumes—with richer qualitative data that better captures the market's mood. This combo of qualitative and quantitative evidence tends to strengthen the forecasting accuracy of inventory prediction algorithms, specially in fast-moving sectors wherein investor sentiment drives buying and selling behaviour.

The predictive capacity of information-derived sentiment has been demonstrated across a number of investigations, which have hired more than a few machine-mastering and deep-mastering techniques. for example, Khonde and associates (2024) combined both vector-area models and lexicon-primarily based classifiers with lengthy quick-time period reminiscence (LSTM) networks to categorise Malaysian monetary headlines as tremendous or bad. Their approach yielded an 86-in keeping with-cent accuracy charge for sentiment tagging and 80 three in line with cent at the same time as forecasting the direction of inventory actions that followed, thereby highlighting how efficaciously news sentiment can beautify conventional time-collection prediction frameworks.

latest improvements in forecasting techniques, significantly the Context-aware Bidirectional long brief-term reminiscence version—abbreviated as CAB-LSTM—have achieved noteworthy gains in accuracy via weaving collectively topic modelling with sentiment scores derived from monetary headlines. In a practical take a look at, Xiao (2024) determined that CAB-LSTM always outperformed trendy forecasting frameworks whilst deployed on news tales about Malaysian stocks, correctly forecasting each sectoral tendencies and price shifts. The model's dual capacity to sense contextual relevance while gauging the intensity of sentiment proved pivotal for securing better directional forecasts, especially during turbulent market phases.

Within the Malaysian financial landscape, sentiment has emerged as a powerful predictor both at the sectoral and the broader macroeconomic levels. Chan and colleagues (2024) documented long-lived, asymmetric impacts of sentiment on different stock prices, with especially pronounced effects in the finance and consumer sectors. Paradoxically, the technology sector demonstrated a marked insensitivity to these sentiment swings, implying that investor psychology varies significantly from one industry to the next. helping this broader view, Ho and Ong (n.d.) illustrated how sentiment alerts lifted from newspapers can nowcast Malaysian commercial enterprise self belief and forecast the trajectory of private investment growth with a lead time of two to three quarters. Their findings thus extend the relevance of sentiment analysis well beyond the confines of equity markets.

Transformer architectures have steadily gained popularity not only for their efficacy with text but also for the seamless way they can accommodate numerical inputs. In a recent study, Dong and colleagues (2024) examined how sentiment scores from social media and news comments could be blended with traditional Transformer frameworks to forecast stock price movements. They reported that this dual-stream input—sentiment data running in parallel with standard technical indicators—produced both sharper prediction accuracy and a noticeably shorter latency between data arrival and output. Their results were consistent with earlier work by Chou and Ramachandran (2021), which demonstrated that coupling time-series models with extracted news sentiment substantially cut forecast errors and made markets' responses quicker to unfolding events.

The larger picture of collective data usage has also evolved, thanks to newly minted datasets designed specifically for the financial domain. The FNSPID set conceived by Dong et al. (2024) stands out in this regard; it pairs minute-by-minute stock prices with timestamped sentiment tags, thereby granting researchers a rich, synchronized resource for training sentiment-enhanced predictive networks. Building on this momentum, Kurisinkel and co-authors (2024) underscored the value of event-driven language models that feed in fresh news as it breaks, allowing forecasts to refresh themselves in near-real time. Such an architecture points toward a markedly more flexible blueprint for financial prediction, one that aims to keep pace with the constant ebb and flow of market-relevant information.

While significant headway has been made in applying sentiment analysis to financial markets, obstacles remain. A particular hurdle in the Malaysian context is the limited availability of well-annotated sentiment datasets that capture local nuances. Without a robust pool of quality labels, there is little room to effectively train or rigorously test forecasting models. adding to this project, the predictive strength of sentiment scores frequently hinges on several external factors: the reliability of the news outlet, the specifics of the information occasion itself, and the period among publication and observable market movement. these variables signal a urgent need for studies this is both extra geographically focused and methodologically bendy.

## 2.5 Deep Learning for Stock Price Prediction

### 2.5.1 Long Short-Term Memory (LSTM)

Long-Short Term Memory (LSTM) models have emerged as a fave within the finance literature. Not like traditional forecasting methods—be it ARIMA, linear regression, or transferring averages—LSTMs are not shackled with the aid of assumptions of linearity or strict stationarity. Their architecture is designed specially to capture complicated, nonlinear relationships over prolonged time frames, making them well acceptable for predicting inventory fees. The technical edge of LSTMs lies of their gating gadgets, which carefully manage the passage of records at some point of the sequence. by doing so, they in large part overcome the vanishing gradient issue that regularly hinders preferred recurrent neural networks. As a end result, LSTM fashions can keep crucial past indicators alive whilst filtering out noise, aligning them intently with the unpredictable rhythms regular of financial markets (Gaurav et al., 2023; Zhong, 2024).

A developing frame of empirical studies highlights the effectiveness of lengthy quick-time period reminiscence (LSTM) networks for forecasting stock prices. Pan (2024) lately showed that an LSTM structure without a doubt outperformed aid vector regression, choice timber, and linear regression models whilst tested on excessive-profile equities like Apple and Tesla; the LSTM performed each a decrease imply squared errors and a higher coefficient of determination. In a exceedingly different environment, Ku et al. (2023) carried out the same approach to the Malaysian inventory market and enriched the version inputs with investor insights and numerous technical indicators. This integration now not handiest boosted prediction accuracy however additionally led to significantly better cumulative returns than those yielded by means of random choices or conventional choice strategies.

Researchers have additionally superior LSTM frameworks by means of experimenting with feature engineering and hybrid architectures. Ozupek et al. (2024) introduced a model that pairs LSTM with empirical mode decomposition and several broadly used technical indicators, yielding importantly more potent consequences for fashion forecasting. inside the area of commodity expenses, Brown et al. (2024) stated that a similar LSTM configuration produced a lower suggest absolute errors in oil-rate predictions when as compared against traditional time-series methods which include ARIMA and exponential smoothing.

LSTM certainly provide powerful predictive skills, yet they arrive with continual drawbacks. training those architectures commonly needs sizable computational strength together with large, datasets, each of which may be in quick supply, particularly for smaller economic groups. Furthermore, LSTMs behave like so-called "black boxes," meaning that analysts and stakeholders often struggle to understand precisely how the models arrive at their forecasts—an interpretive gap that is particularly troubling in the risk-sensitive world of finance (Alamu & Siam, 2024). Encouragingly, ongoing research into explainable artificial intelligence and various model-tuning techniques is steadily enhancing both the transparency and the overall effectiveness of LSTM-based systems for financial prediction.

### 2.5.2 Gated Recurrent Unit (GRU)

To address some of these challenges, the Gated Recurrent Unit (GRU) has emerged as a popular alternative for time-series tasks in the financial domain. Proposed by Chung and colleagues in 2014, the GRU reduces the architecture's computational burden by folding the separate forget and input gates into a single update gate, which in turn eliminates the need for dedicated memory cells. This streamlined design allows GRUs to be trained faster and to make predictions with lower resource overhead compared to full-scale LSTMs. As a result, the GRU is often the desired choice in scenarios stressful rapid inference, along with actual-time stock fee prediction on the unexpectedly converting Malaysian marketplace.

current proof factors to the developing effectiveness of Gated Recurrent Unit (GRU) fashions inside the field of monetary forecasting. In his 2024 evaluation, Makinde reviews that a GRU architecture, first-rate-tuned with the Adam optimization set of rules, produced a lower Root mean square error (RMSE) even as converging more quick than traditional LSTM networks while carried out to day by day stock-charge predictions in rising markets. Kalbaliyev and Szegedi (2020) corroborate those findings by noting that the GRU framework yielded superior overall performance all through extended prediction horizons, producing large buying and selling returns and smaller mistakes margins, especially whilst marketplace situations have been erratic.

Such flexibility to house the moving, nonlinear characteristics of monetary time series makes the GRU specifically appropriate for modeling Malaysian stock indices. Chen et al. (2023) recently introduced a GRU variant that integrates multi-stock functions from distinctive sectors, demonstrating that this move-enterprise facts fusion method curbs overfitting and extensively enhances model generalization. This approach is well timed for Malaysia, in which interdependencies amongst sectors can quick adjust basic market sentiment.

Hybrid models that pair GRU with other architectures have additionally started to expose promise. A 2022 take a look at titled “stock fee Prediction the use of Bi-LSTM and GRU-based Hybrid Deep gaining knowledge of technique” located that fusing the bidirectional reminiscence of Bi-LSTM with the computational performance of GRU caused higher lengthy-time period trend forecasts than either model may want to achieve in isolation across numerous stock datasets.

latest traits have demonstrated that gating mechanisms like those in gated recurrent unit (GRU) can paintings distinctly well while nested inside ensemble architectures. of their 2024 examine, He and colleagues reported on a filtering ensemble that harnesses numerous parallel GRU networks; the innovation has already translated into constant profits in forecast precision over medium to lengthy horizons. perhaps most putting, the ensemble exhibits stable performance regardless of the precise time frame concerned, a functionality that investors within the Malaysian market have located beneficial given that volatility frequently ripples outward from worldwide and nearby occasions.

but the generation isn't a common panacea. Velarde and co-authors mentioned in 2022 that GRU profits tend to flatten out whilst carried out to series displaying little temporal range; in such cases, classical approaches, along with easy moving averages, can on occasion identical or surpass state-of-the-art architectures. The researchers additionally cited that bringing in sentiment ratings derived from financial journalism and social media chatter can help close the performance hole that also lingers among the GRU and LSTM community. Their findings underscore a broader lesson: for effective forecasting in finance, depending totally on one sort of statistics or one modeling paradigm is rarely sufficient.

### 2.5.3 Attention-Based CNN-LSTM (ACNN-LSTM)

ACNN-LSTM architectures have begun to draw interest in the discipline of economic forecasting, mainly for stock rate predictions that draw on both textual and numerical inputs. by pairing the neighbourhood characteristic extraction power of Convolutional Neural Networks with the sequence-mastering prowess of long brief-term reminiscence units, and then augmenting the aggregate with an attention layer that highlights the maximum salient indicators, those hybrid models are capable of expand forecasts that experience extra context-sensitive. The practice of feeding sentiment-rich information headlines into the gadget along uncooked price information is a major reason why the method has shown promise.

In the first stage, CNN blocks process structured time-series information—such as price charts or fixed-length vectors obtained from embedded text—to mine for short-term trends and recurrent seasonal motifs. Next, the LSTM component steps in to unravel longer temporal dependencies, shedding light on how previously detected patterns drift and change. Attention modules, like the Convolutional Block Attention

Module (CBAM), further streamline the workflow by automatically assigning higher weights to features that have proven useful in the past. In a recent study, Li et al. (2023) integrated CBAM within their CNN-LSTM pipeline for BBVA stock data and reported an impressive Mean Absolute Error of just 0.0058 and an $R^2$ score of 0.9673, marks that considerably eclipse those of more conventional CNN-LSTM configurations.

Recent research on hybrid architectures has turned to combining convolutional long short-term memory (CNN-LSTM) networks with gradient-boosted tree methods and attention layers in order to leverage the character benefits of each approach. Zhu and Chen (2023) constructed a machine wherein function representations extracted through a CNN-LSTM spine have been fed into an severe gradient boosting (XGBoost) regressor. Their effects showed that this -level pipeline completed better accuracy than either the neural network alone or the tree model used separately. In a related study, Shi et al. (2022) embedded an additive attention mechanism directly into their CNN-LSTM framework, reporting gains in directional prediction performance for stock price movements.

Zhang et al. (2023) undertook a comparative analysis involving a convolutional bidirectional LSTM network augmented with attention, pitting it against both standard LSTM and CNN-LSTM models across multiple financial datasets. The attention-enhanced architecture repeatedly outperformed its peers, confirming the utility of more nuanced weightings on past inputs. Wu et al. (2024) added further evidence by evaluating a CNN-LSTM variant outfitted with what they termed an additive attention mechanism (AAM); their findings indicated an 8 to 10 percent reduction in forecast error compared to the baseline CNN-LSTM, thus highlighting the dual benefits of increased accuracy and stability.

Incorporating unstructured text data—such as news headlines—into ACNN-LSTM frameworks substantially expands their potential to gauge investor sentiment. Li and Hua (2024) showed that through embedding attention modules, the structure is capable of highlight phrases wealthy with sentiment. whilst these highlighted cues are integrated with traditional stock indicators and technical metrics, the model's generalization performance on Malaysian forecasts markedly improves. Similarly, Liu and Zhang (2022) argue the attention mechanism augments ACNN-LSTM's resilience to market volatility by allowing the system to pivot toward the most relevant financial signals that mirror prevailing sentiment.

The seamless integration of localized feature extraction, temporal processing, and real-time attention lets ACNN-LSTM exploit diverse data streams, thereby lifting forecasting accuracy. More importantly, this attention layer bolsters the model's interpretability, clearly indicating which features or time steps attract focus. Such transparency is especially critical in finance, where stakeholders demand robust explanations alongside numerical predictions.

### 2.6 Previous Work on Sentiment Analysis in Financial Forecasting

Sentiment evaluation has steadily received prominence in inventory forecasting partly due to the fact social media and on line news now flood the general public with opinions. Academics and market analysts are testing numerous strategies to funnel the emotional undercurrents from forums, tweets, press releases, and conference calls into algorithmic forecasts, hoping this will yield a sharper snapshot of market mood.

latest research suggests that integrating sentiment analysis with time-series inventory records can tremendously improve predictive accuracy. as an example, neural architectures which include convolutional neural networks and long short-term reminiscence groups, when fed sentiment indicators extracted from on line inventory boards, have outperformed traditional forecasting techniques (Jing et al., 2021). similarly, hybrid fashions that pair the VADER sentiment lexicon with LSTM frameworks using social media streams have yielded superior results (Dutta et al., 2021). Sulistianingsih and Martono (2024) extended these approaches to the Indonesian market, thereby demonstrating that sentiment-driven techniques retain their effectiveness across diverse linguistic and cultural settings.

On a more advanced front, transformer-based architectures like BERT have been employed to mine sentiment from news headlines and corporate earnings transcripts (Fu & Zhang, 2024; Sarkar & Shahid, 2025). The deep contextual embeddings generated with the resource of these models permit for a extra nuanced interpretation of language, translating into giant upgrades in predictive energy. but, practitioners even though deal with limitations such as casual wording, sarcasm, and the general noisiness of short-text data, challenges that underscore the necessity of thorough preprocessing and resilient language models.

Overall, the reviewed literature supports the integration of sentiment analysis in financial forecasting models, reinforcing its relevance to this study's aim of improving stock price prediction through deep learning and news sentiment.

Table 2.1 summarizes key studies that have applied sentiment analysis techniques in financial forecasting.

| Author/Year | Title | Research Focus | Techniques Used |
|---|---|---|---|
| Jing et al. (2021) | Stock Forum Sentiment Analysis Using CNN for Price Movement Forecast | Shows domain-specific sentiment improves forecasting accuracy | CNN-based sentiment classification |
| Dutta et al. (2021) | Sentiment Analysis Using VADER and LSTM for Market Behavior Prediction | Hybrid sentiment + LSTM enhances market movement prediction | VADER + LSTM hybrid model |
| Khonde et al. (2024) | Forecasting Stock Trends with LSTM and News Sentiment Analysis | Validates value of integrating sentiment with sequential models | Sentiment classification + LSTM |
| Sulistianingsih & Martono (2024) | LSTM-CNN Hybrid with Lexicon Scoring for Indonesian Stock Forecast | Combines emotional tone and time-series for more stable predictions | Lexicon-based scoring + LSTM-CNN |

| | | | |
|---|---|---|---|
| Zhu et al. (2023) | Attention-Based CNN-LSTM-XGBoost for Stock Price Forecasting | Leverages attention to improve nonlinear sequence learning for price movement | Attention-based CNN-LSTM + XGBoost |
| Fu & Zhang (2024) | BERT–LLA for Sentiment and Technical Indicator Fusion | Demonstrates strong performance in combining text and numeric trends for forecasting | BERT–LLA (BERT + Lookahead Attention) |

**Table 2-1 Previous Work Study**

## 2.7 Research Gap

While a lot of research has demonstrated the effectiveness of sentiment-based forecasting models in Western financial markets, there is a notable gap in the application of such methods in Southeast Asian contexts. Specifically, the Malaysian banking sector remains underexplored in this regard. Most existing studies either focus on global indices or rely solely on historical numerical data, overlooking the predictive potential of localized sentiment data.

Moreover, limited research has compared multiple deep learning models (LSTM, GRU, ACNN-LSTM) within a unified framework applied to Malaysian banks. This study aims to bridge this gap by analyzing the predictive relationship between news sentiment and stock prices of CIMB and Maybank, offering a comprehensive and context-specific contribution to the literature.