

Image-Driven Stock Price Prediction with LLaMA: A Prompt-Based Approach

Qizhao Chen

Graduate School of Information Science, University of Hyogo, Kobe, Japan

Email: af24o008@guh.u-hyogo.ac.jp

Manuscript received February 26, 2025; accepted April 2, 2025; published April 17, 2025

Abstract—This study proposes a novel framework for predicting stock price movements using a prompt-based approach with the Large Language Model Meta AI (LLaMA) model, where candlestick charts serve as the primary input. Unlike traditional deep learning models that process images through convolutional or transformer-based architectures, the proposed method leverages LLaMA’s prompt-driven reasoning to interpret financial chart patterns. In addition, a teacher-student model incorporating LLaMA and Qwen is explored. To assess the effectiveness of this prompt-based Large Language Model (LLM) approach, its performance is compared with established models, including Convolutional Neural Network (CNN), ResNet, and Vision Transformer. Experimental results demonstrate that the proposed method consistently outperforms these deep learning models, highlighting the potential of prompt-based LLM techniques for financial time series forecasting using visual inputs.

Keywords—Convolutional Neural Network (CNN), Large Language Model Meta AI (LLaMA), Residual Neural Network (ResNet), vision transformer, stock price prediction

I. INTRODUCTION

The stock market is a dynamic system influenced by various factors, such as technical indicators, macroeconomic data, and market sentiment. Among these factors, technical indicators, derived from historical stock and volume data, are commonly used in stock price prediction [1–5]. Typically, they are input as numerical values into machine learning models to forecast stock prices or movements. Although the numerical representation of technical indicators provides valuable information, they may not fully capture the intricate patterns and trends embedded in the data [6]. To address this problem, traders and investors often visualize stock and volume data using stock charts such as candlestick charts. Traditionally, traders and investors manually analyze these charts to inform their trading decisions, which is time-consuming and subjective. With the rapid development of deep learning, automated pattern recognition has become increasingly feasible. Models like Convolutional Neural Networks (CNNs), Residual Neural Networks (ResNets), and Vision Transformer (ViTs) excel at extracting features from visual data and identifying patterns that humans might miss. These advances motivate the exploration of image-based approaches to stock price prediction [7, 8].

However, using an image-based approach for stock price prediction poses several challenges. For instance, stock charts often contain large areas of non-informative space, which makes it difficult for the model to focus specifically on the sections that reflect stock movements. Furthermore, identifying global patterns from local patterns within sub-charts can be challenging. This challenge arises because sub-charts often highlight short-term fluctuations or isolated

trends, which obscures the broader trends or relationships necessary for accurate stock price prediction. Creating a model that can effectively integrate and interpret these localized insights to capture the broader context is a complex yet crucial task [9].

Based on experimental findings and previous research, images provide a more effective way for predicting stock trends, especially over long periods, compared to numerical data. This suggests that the image-based approach is more suited to long-term investment strategies compared to high-frequency trading, which relies heavily on quick decisions based on real-time numerical data.

Recent advances in Large Language Models (LLMs), such as Large Language Model Meta AI (LLaMA), have demonstrated remarkable capabilities in reasoning, pattern recognition, and multimodal understanding. Although traditionally designed for natural language processing, LLMs can be adapted to analyze visual inputs through prompt-based methods. One of the examples is the Llama 3.2-Vision instruction-tuned model. It is designed and optimized for visual recognition, image reasoning, captioning, and answering general questions about an image. Furthermore, unlike Convolutional Neural Network (CNN), ResNet, and ViT, which directly extract features from images, a prompt-driven approach takes advantage of the pre-trained knowledge and reasoning ability to infer stock market trends from candlestick chart representations.

Recently, some researchers have used prompt-based methods with proposed LLMs such as finMA [10] and finGPT [11] for stock price forecasting. However, these approaches often require manual input of extensive data into the prompt, including macroeconomic indicators, financial news, and historical stock data. This manual input requires significant extra effort in data collection and search. Therefore, a more streamlined, yet effective, prompt-based approach could be a better alternative.

Building on the advancements of generative LLM, this study proposes a novel framework that utilizes LLaMA with a prompt-based approach to analyze candlestick charts and predict stock movements. By converting financial images into structured prompts, the model can generate insights that go beyond traditional deep learning techniques, offering a new perspective on image-based stock price forecasting.

The main contributions of this research are as follows.

- 1) This research proposes a novel framework for stock price prediction that leverages stock images as input in a prompt-based approach using the LLaMA model. By transforming financial charts into structured prompts, the proposed method enables LLaMA to analyze intricate market patterns and trends, offering a new

perspective on image-based financial forecasting. In addition, a teacher-student model with LLaMA and Qwen is also explored. The results show that this approach can further improve the prediction performance.

- 2) The comparative analysis between prompt-based LLaMA approach and traditional deep learning models such as CNN, ResNet, and ViT is conducted. The results show that the proposed method improves automated pattern recognition and reduces the reliance on manual feature extraction.

The remainder of this paper is structured as follows. Firstly, related work is listed in Section II. The methodology is then described in Section III. In Section IV, the experimental results are presented. Finally, we conclude with Section V.

II. RELATED WORK

A. Numerical Feature-Based Models

Traditional machine learning models such as Support Vector Machine (SVM), Random Forest, and Gradient Boosting have been widely used in stock price forecasting. The input of all these methods is numerical features derived from historical prices, volume, and textual data. For example, Hu *et al.* [12] leveraged SVM to predict the stock price with four company-specific and six macroeconomic factors. The results demonstrate that SVM is a robust tool for stock price prediction. Zheng *et al.* [13] employed a Random Forest model to predict stock price trends, demonstrating its effectiveness in forecasting stock movements with high accuracy. Their studies show that Random Forest can outperform models like SVM and logistic regression, particularly for long-term predictions. Nakagawa and Yoshida [14] proposed a time-series gradient boosting tree model that integrates time-series and cross-sectional data for stock price prediction. Their method uses Dynamic Time Warping (DTW) to capture temporal patterns and demonstrates superior predictive accuracy and profitability compared to previous approaches, particularly in forecasting monthly stock price changes.

B. Deep Learning for Sequential Data

To handle sequential data, Recurrent Neural Network (RNN) such as Long Short Term Memory (LSTM) has demonstrated outstanding performance in modeling temporal dependencies in time-series data. In addition, Transformer-based models, which utilize self-attention mechanisms, have further improved the performance by capturing long-term and dynamic dependencies. Many studies have shown the superior performance of deep learning models over traditional models such as SVM. Ancy *et al.* [15] used a hybrid RNN model to predict the trend of the stock price with stock features and financial news, and the results show that the proposed hybrid RNN model with the sentiment of the news as input outperforms all other indicators. Abdelfattah *et al.* [16] used sentiment data and historical stock market data to improve the prediction of stock market movement and the results show that LSTM has better performance compared to traditional methods such as Neural Network, SVM and Naïve Bayes. Chen and Kawashima [17] compared the performance of several machine learning and

deep learning models in the prediction of stock prices with technical indicators and sentiment data as features and find that Informer, an advanced transformer-based model, can outperform all other models, such as LSTM, SVR and Random Forest in most cases.

C. Image-Based Models

Although models such as CNN have been widely used in various image recognition tasks, they remain underexplored in the context of image-based stock price prediction. Recent studies have begun to address this gap. For example, Bang and Ryu [18] utilized CNN to predict stock movements using candlestick charts as input. Although predictive performance is limited, their study highlights the potential of employing CNNs to forecast stock prices through visual representations of stock data and technical indicators.

Furthermore, pre-trained architectures like ResNet and ViT, renowned for their state-of-the-art performance across numerous computer vision tasks, offer promising alternatives for feature extraction from financial chart images. Some research has begun to explore this avenue. For example, Rosdyana *et al.* [19] used CNN, ResNet, and the VGG network to predict the stock movement in the next day with candlestick charts. Their approach achieves impressive accuracy rates of 92.2% and 92.1% for the Taiwan and Indonesian stock market datasets, respectively. Hung and Chen [9] proposed a Deep Predictor for Price Movement (DPP) model to predict stock movements using candlestick charts. Compared to baseline models such as Prophet, IEM, and LSTM, DPP demonstrates superior performance, particularly in scenarios involving frequent price fluctuations. In addition, Li *et al.* [20] introduced a novel perspective by utilizing pretrained vision transformers for classification of line graph images. The results highlight the effectiveness of vision transformers in handling images.

These developments suggest that the use of such advanced architectures can unlock new opportunities in image-based financial modeling.

III. METHODOLOGY

A. Data Collection and Data Transformation

The daily stock price data of Apple (AAPL), HSBC (HSBA.L), Pepsi (PEP), Tencent (0700.HK), and Toyota (7203.T) from 01/05/2015 to 08/05/2024 are downloaded from Yahoo Finance using the Python library `yfinance`. The downloaded data for each stock contains columns “Open”, “Close”, “High”, and “Low”.

Two types of images are tested in this paper: candlestick charts [9] and bar-chart images [18], as they have been used in previous research. The models evaluated include CNN, ResNet, and ViT. Compared to candlestick charts, bar-chart images are less computationally expensive for models to process.

Fig. 1 shows the overall design of the experiment.

- 1) Candlestick Chart: Candlestick charts can be converted from stock prices using the Python library `mplfinance`. To enhance the image features before feeding them into the LLM, several pre-processing techniques and feature engineering steps are applied. For example, volume data are added below the candlestick chart to capture

trading activity. Fig. 2 is an example of the candlestick chart.

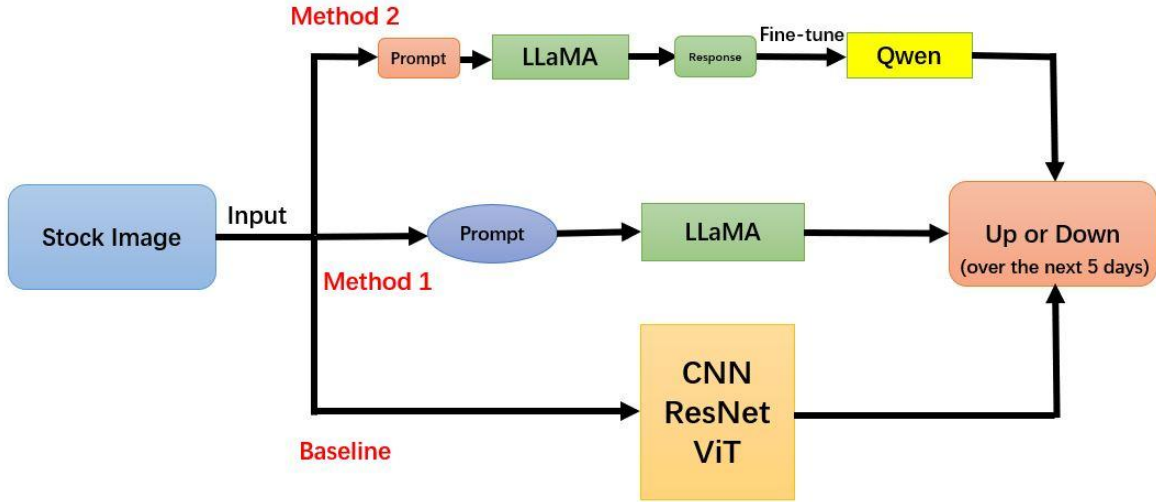


Fig. 1. Whole picture of methodology.

2) Bar-Chart Image: For conversion of bar-chart images, daily data for each stock are normalized to the range [0, 255], which allows prices to be mapped to the pixel intensities. For each trading day, three channels are constructed to represent the open price, the range between the low and high prices, and the close price. The open and close prices are mapped to individual pixel, while the low to high range is represented by a vertical line. These representations are combined into images with dimensions of $256 \times (3 \times \text{input window})$, where 256 corresponds to the normalized price range, and the width reflects the input window size. This process produces structured and image-based representations of stock price movements. Fig. 3 is an example of the bar-chart image.

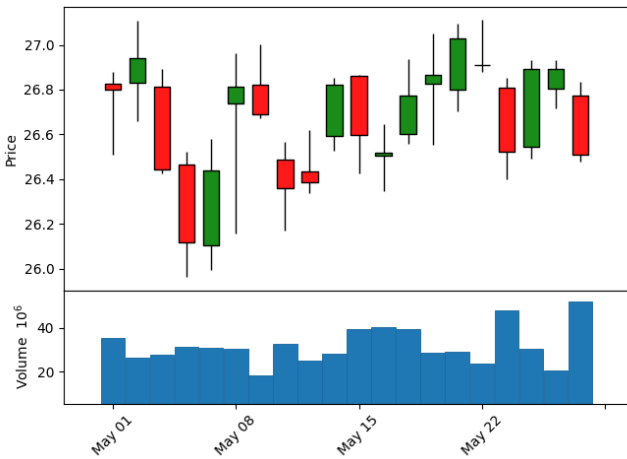


Fig. 2. Candlestick chart input sample.

B. Dataset Preparation

The dataset used in CNN, ResNet and ViT is divided into training and testing sets, with a split of 80:20. Training data include the first 80% of the time series, while the remaining 20% is reserved for testing. Binary labels are generated to indicate whether the price of the stock increases (1) or decreases (0) over the next five trading days. For the LLaMA model, only the candlestick charts with a sliding window size

of 20 days are used as input. This is primarily based on the experimental results of CNN, ResNet, and ViT, which will be presented in Section IV. The results show that models achieve better performance when using candlestick charts. In addition, the sliding window size does not significantly affect model performance.



Fig. 3. Bar-chart Image sample.

C. LLM Model and Prompt Design

LLaMA is a family of advanced language models developed by Meta, designed to perform various natural language processing tasks with high efficiency. Unlike traditional AI models, LLaMA is optimized for reasoning, contextual understanding, and adaptability across different domains. Despite being primarily trained for text-based applications, its capabilities can be extended through prompt engineering. This technique enables the model to process non-textual data, such as structured image inputs.

The model version used in this research is “Llama-3.2-11B-Vision-Instruct”, which is optimized for visual recognition and image reasoning.

The student model used in Method 2 (Table 1) is “Qwen2-VL-7B-Instruct”. Qwen is a series of large language models developed by Alibaba Cloud. “Qwen2-VL” achieves state-of-the-art performance on visual understanding benchmarks, such as MathVista, DocVQA, RealWorldQA, and MTVQA.

Table 1. Prompt for stock price prediction

Method	Prompt	Description
1	<p>Analyze the provided 20-day candlestick chart of company name and predict the stock price movement within the next five days.</p> <p>Key Analysis Points:</p> <ul style="list-style-type: none"> • Technical Indicators: Assess trends using SMA, EMA, RSI, MACD, Bollinger Bands, and volume changes. Highlight any crossovers, divergences, or extreme values. • Candlestick Patterns: Identify bullish or bearish patterns (e.g., engulfing, doji, hammer, shooting star) and explain their significance. • Support/Resistance Levels: Pinpoint recent highs, lows, and key price levels acting as support or resistance. • Market Sentiment: Consider overall sentiment from recent news, earnings reports, or macroeconomic events that could influence the stock price. <p>Output Format (JSON):</p> <pre>{ "prediction": "Up" "Down" "Same", "probabilities": { "Up": 0.XX, "Down": 0.XX, "Same": 0.XX }, "justification": "Technical indicators (e.g., RSI at 70 indicates overbought), patterns (bullish engulfing), support / resistance levels, and market sentiment (Positive due to strong earnings report)."</pre>	<p>This prompt incorporates the technical indicators, market sentiment, and a confidence score for a more comprehensive evaluation. It provides a longer-term perspective by predicting stock movement over the next five days, offering deeper insights into trends and potential price shifts.</p>
2	<p>You are a financial analyst who predicts stock prices. Analyze the candlestick chart of company name and predict the stock price movement over the next five trading days. Include: 1. Technical Indicators (e.g., SMA, EMA, RSI, MACD, Bollinger Bands, volume). 2. Candlestick Patterns (e.g., engulfing, doji, hammer, shooting star). 3. Support/Resistance Levels (recent highs/lows). 4. Market Sentiment (e.g., news, external factors). Weigh signals to determine likelihood of Up, Down, or Same. Response Format: 1. Prediction: Up/Down/Same. 2. Probabilities for each scenario. 3. Justification based on indicators, patterns, sentiment.</p>	<p>This is the prompt used by Qwen, which will be fine-tuned using the output from the LLaMA model.</p>

Model fine-tuning is conducted using Unsloth¹, which is an optimization library that enhances the efficiency of fine-tuning large language models. It supports memory-efficient techniques like LoRA, which enables faster training on consumer GPUs with reduced resource consumption. In this study, Unsloth is applied to fine-tune Qwen2-VL-7B-Instruct for improved visual recognition and image reasoning. For fine-tuning, the dataset will be structured according to the format shown in Table 2.

To improve forecast performance, two different approaches are examined. The methods used in this paper are given in Table 1.

In Method 1, the Llama-3.2-11B model is used directly to predict stock movement without any fine-tuning. Method 2 applies the concept of knowledge distillation. Knowledge distillation is a technique in which a smaller model (the student) learns to replicate the behavior of a larger, more complex model (the teacher). In this study, the Llama-3.2-11B model serves as the teacher, generating stock movement predictions and corresponding responses to guide the student model. These responses are then used to fine-tune the Qwen2-7B model. This approach is especially beneficial as it eliminates the need for human-labeled data. In addition, it enables the transfer of knowledge from a large,

high-performing model to a smaller one. As a result, the smaller model can achieve comparable performance while being more computationally efficient and faster. By applying knowledge distillation, the predictive power of Llama-3.2-11B is leveraged, while the capabilities of Qwen2-7B are refined. This approach enhances the accuracy and reliability of stock forecasting.

D. Baseline Model

- 1) Convolutional Neural Networks (CNNs): The architecture of the CNN model used in this paper consists of two convolutional layers, each followed by Leaky ReLU activation and max pooling. The first convolutional layer uses eight filters of size 3×3 with a padding of one, preserving the original image dimensions. The second convolutional layer uses 16 filters of the same size. Max pooling with a 2×2 filter and stride of two are applied after each convolutional layer to reduce the image dimensions progressively. Finally, the extracted features are flattened and passed through a fully connected layer with a softmax output for binary classification.
- 2) Residual Neural Network (ResNet): ResNet, proposed by [21], is a deep learning architecture designed to address the problem of vanishing gradients. By using the skip connection structure, ResNet enables the

¹ <https://unsloth.ai/>

network to learn a residual mapping rather than the full transformation. This makes it easier for the network to optimize.

In this research, the ResNet-18 architecture is used. ResNet-18 consists of one initial convolutional layer, sixteen convolutional layers within the residual blocks, and one fully connected layer.

When training the ResNet-18 model, bar-chart images are resized to 224×224 pixels to meet the model's fixed input size requirements. The images are then normalized to the range $[-1, 1]$ using standard mean and standard deviation values. This normalization is necessary because ResNet-18 uses pre-trained weights that assume the input pixels are already normalized. In addition, each image is converted to a three-channel format to match the design of ResNet, which expects RGB images with three color channels (Red, Green, and Blue).

3) Vision Transformer (ViT): ViT, introduced by [22], is a type of encoder transformer designed for computer vision. ViT takes the images as input by decomposing them into fixed-size patches and processing them as a sequence using self-attention mechanisms. This architecture can capture global contextual information more effectively than CNN.

In this research, the vit_b_16 architecture from the torchvision library is used. It is the same model introduced in the original paper: *An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale* [22]. The images are divided into non-overlapping patches of size 16×16 pixels. These patches are treated as input tokens for the transformer.

E. Evaluation

The models are evaluated using analysis of accuracy, precision, recall and F1-Score. Accuracy measures the percentage of correct predictions made by the models. Precision can be used to show how many of the positive predictions made by the models are actually positive, and this metric focuses on positive predictions. Recall, on the other hand, indicates the proportion of all actual positive cases that are correctly identified by the model. The F1-Score is a combination of Precision and Recall and provides a balanced measure.

Since LLaMA outputs stock movement predictions with probabilities, the movement with a probability exceeding 50% will be the model's final decision. For example, if the model predicts an "Up" movement with a 70% probability, the final decision is that the stock is expected to rise over the next five days.

IV. RESULTS

The results of the stock movement forecasting are shown from Table 4 to 5. Table 3 shows a sample response from the LLaMA model.

The results in Table 4 demonstrate that using LLaMA for the prediction of stock movements with candlestick charts yields strong performance, particularly when Method 2, a teacher-student model approach, is applied. Across the five stocks, applying Method 2 consistently improves accuracy, precision, recall, and the F1-Score. In particular, Toyota

benefits the most from Method 2, achieving a precision of 0.90 and an F1-Score of 0.93, the highest among all the stocks tested. Apple, HSBC, Pepsi and Tencent also show improvements. For example, Apple achieves an accuracy of 0.90, while Tencent reaches an accuracy of 0.88 after fine-tuning. The improvements are particularly visible in accuracy and precision, while the recall remains relatively stable. These results suggest that LLaMA achieves relatively strong performance using only its pretrained knowledge. However, after knowledge distillation, the model gains improved knowledge. This enhancement further boosts its predictive capabilities.

Table 2. Dataset format

```
[
{"role": "user",
"content": [{"type": "text", "text": "prompt", "type": "image",
"image": candlestick chart}, {"role": "assistant",
"content": [{"type": "text", "text": "answer"}]}
]
```

Although these results highlight the predictive power of LLaMA, it is equally important to consider the model's interpretability in financial decision making. As shown in the sample response from the LLaMA model (Table 3), LLaMA forecasts stock movements by analyzing candlestick chart patterns, technical indicators, and market sentiment, which are the main factors affecting the stock movement. Specifically, the structured prompt guides LLaMA to analyze technical indicators like SMA, EMA, and MACD. It is then tasked with recognizing common candlestick patterns, such as doji, engulfing, and hammer formations. These patterns are then integrated with historical data. In addition, LLaMA assesses market sentiment from recent news articles, macroeconomic events, and financial documents by leveraging its pre-trained knowledge. By extracting meaningful insights from charts and correlating them with stock price behaviors,

LLaMA's predictions are not only accurate but also interpretable.

Besides, when comparing the LLMs with CNN, ResNet, and ViT (Tables 5 and 6), LLaMA and Qwen generally have better performance. These results suggest that LLMs have the potential to outperform deep learning models in stock image analysis and stock movement forecasting.

For CNN, ResNet and ViT, the results (Tables 5 and 6) indicate that different deep learning models exhibit varying levels of effectiveness when using stock images to forecast stock movement. CNN and ResNet generally outperform the Vision Transformer model when using candlestick charts. The highest performance is observed when using candlestick chart images, where CNN and ResNet achieve strong accuracy, precision, and recall across all stocks and sliding window sizes. In contrast, ViT consistently lags behind, showing lower accuracy and weaker F1-Scores. The performance of the models shows little difference when using different sliding window sizes, which suggests that when using images as input, sliding window size does not affect the performance.

Table 3. Response sample from LLaMA

<p>Step 1: Analysis of Technical Indicators The SMA and EMA are trending upwards, indicating a bullish trend. The RSI is above 60, but not at extremely high levels, suggesting the stock is slightly overbought but not at a point of caution. The MACD is also bullish, with the line above the signal line and moving upwards, supporting the upward trend. Bollinger Bands are wide, but price is above the middle band, confirming the uptrend. Volume has been increasing, but not dramatically.</p>
<p>Step 2: Candlestick Patterns The chart shows a series of bullish candles, including a hammer pattern on May 15, indicating a potential reversal of the downward trend and a rise in the next day. The overall trend since May 15 is upward.</p>
<p>Step 3: Support/Resistance Levels Recent highs (around \$29.5) and recent lows (around \$28.2) provide support and resistance levels. The stock price closed at around \$29.0 on the last day, above the middle of the range.</p>
<p>Step 4: Market Sentiment No recent news or macroeconomic events strongly influencing the stock price are noted. However, a strong earnings report from Apple could continue to drive the stock price upwards.</p>
<p>Conclusion Based on technical indicators, bullish patterns, support/resistance levels, and market sentiment, the prediction for AAPL stock price movement within the next five days is likely to be up.</p>
<p>JSON Output:</p> <pre>{ "prediction": "Up", "probabilities": {"Up": 0.80, "Down": 0.10, "Same": 0.10 }, "justification": "Technical indicators, bullish patterns, support/resistance levels, and market sentiment due to positive earnings report." }</pre>

Table 4. Prediction performance of LLaMA, candlestick chart

Stock	Scenario	Method	Accuracy	Precision	Recall	F1-Score
Apple	w/o Fine-Tuning	1	0.86	0.91	0.93	0.92
	with Fine-Tuning	2	0.90	0.92	0.93	0.92
HSBC	w/o Fine-Tuning	1	0.84	0.88	0.93	0.90
	with Fine-Tuning	2	0.88	0.89	0.94	0.91
Pepsi	w/o Fine-Tuning	1	0.83	0.87	0.94	0.90
	with Fine-Tuning	2	0.89	0.89	0.96	0.92
Tencent	w/o Fine-Tuning	1	0.81	0.87	0.92	0.89
	with Fine-Tuning	2	0.88	0.89	0.93	0.91
Toyota	w/o Fine-Tuning	1	0.84	0.87	0.95	0.91
	with Fine-Tuning	2	0.92	0.90	0.97	0.93

Table 5. Prediction results for CNN, ResNet and ViT, bart-chart image

Stock	Sliding Window	CNN				ResNet				Vision Transformer			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
Apple	5 Days	0.56	0.70	0.56	0.59	0.61	0.64	0.61	0.62	0.78	0.61	0.78	0.69
	10 Days	0.53	0.69	0.53	0.57	0.77	0.65	0.77	0.69	0.78	0.61	0.78	0.69
	20 Days	0.52	0.67	0.52	0.56	0.76	0.62	0.76	0.68	0.78	0.61	0.78	0.69
HSBC	5 Days	0.69	0.68	0.69	0.68	0.71	0.69	0.71	0.70	0.81	0.65	0.81	0.72
	10 Days	0.73	0.70	0.73	0.71	0.53	0.68	0.53	0.57	0.81	0.65	0.81	0.72
	20 Days	0.75	0.73	0.75	0.74	0.70	0.68	0.70	0.69	0.81	0.65	0.81	0.72
Pepsi	5 Days	0.48	0.75	0.48	0.52	0.79	0.71	0.79	0.72	0.80	0.63	0.80	0.71
	10 Days	0.41	0.74	0.41	0.43	0.79	0.63	0.79	0.70	0.80	0.63	0.80	0.71
	20 Days	0.51	0.73	0.51	0.55	0.79	0.69	0.79	0.71	0.80	0.64	0.80	0.71
Tencent	5 Days	0.65	0.59	0.65	0.61	0.63	0.59	0.63	0.61	0.73	0.53	0.73	0.61
	10 Days	0.64	0.60	0.64	0.62	0.66	0.61	0.66	0.63	0.73	0.53	0.73	0.62
	20 Days	0.66	0.61	0.66	0.63	0.70	0.61	0.70	0.63	0.73	0.54	0.73	0.62
Toyota	5 Days	0.51	0.65	0.51	0.55	0.74	0.67	0.74	0.69	0.77	0.60	0.77	0.68
	10 Days	0.51	0.64	0.51	0.55	0.75	0.67	0.75	0.69	0.77	0.60	0.77	0.68
	20 Days	0.52	0.63	0.52	0.55	0.68	0.63	0.68	0.65	0.77	0.60	0.77	0.67

Table 6. Prediction results for CNN, ResNet and ViT, candlestick chart

Stock	Sliding Window	CNN				ResNet				Vision Transformer			
		Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
Apple	5 Days	0.85	0.85	1.00	0.92	0.85	0.85	1.00	0.92	0.78	0.53	0.52	0.52
	10 Days	0.83	0.83	1.00	0.91	0.77	0.83	0.90	0.87	0.77	0.53	0.51	0.51
	20 Days	0.80	0.80	1.00	0.89	0.75	0.80	0.93	0.86	0.78	0.49	0.50	0.49
HSBC	5 Days	0.82	0.82	1.00	0.90	0.74	0.82	0.88	0.85	0.77	0.51	0.50	0.49
	10 Days	0.83	0.83	1.00	0.91	0.79	0.84	0.93	0.88	0.79	0.45	0.47	0.46
	20 Days	0.83	0.83	1.00	0.91	0.77	0.83	0.90	0.87	0.78	0.50	0.50	0.49
Pepsi	5 Days	0.83	0.83	1.00	0.91	0.74	0.84	0.85	0.85	0.74	0.47	0.49	0.47
	10 Days	0.84	0.84	1.00	0.91	0.81	0.85	0.94	0.89	0.81	0.56	0.53	0.53
	20 Days	0.85	0.85	1.00	0.92	0.81	0.85	0.93	0.89	0.74	0.44	0.47	0.45
Tencent	5 Days	0.82	0.82	1.00	0.90	0.75	0.82	0.87	0.85	0.77	0.50	0.50	0.48
	10 Days	0.86	0.86	1.00	0.92	0.79	0.86	0.90	0.88	0.77	0.53	0.52	0.51
	20 Days	0.85	0.85	1.00	0.92	0.81	0.85	0.94	0.89	0.79	0.46	0.47	0.46
Toyota	5 Days	0.83	0.83	1.00	0.91	0.77	0.82	0.92	0.87	0.79	0.50	0.50	0.50
	10 Days	0.84	0.84	1.00	0.91	0.76	0.84	0.88	0.86	0.78	0.46	0.49	0.46
	20 Days	0.84	0.84	1.00	0.91	0.77	0.84	0.91	0.87	0.77	0.49	0.49	0.49

More specifically, for candlestick charts, both CNN and ResNet demonstrate high accuracy, with the CNN model often exceeding 0.80. This suggests that these two models are particularly effective at identifying stock movement patterns from candlestick chart images. Their strong performance is further reflected in F1-Scores greater than 0.90, indicating a balanced precision and recall. However, ViT struggles significantly with candlestick charts, with its accuracy and F1-Scores remaining around 0.50, which suggests that it cannot capture the crucial visual features needed for accurate predictions.

In contrast, when bar-chart images are used, the overall prediction performance decreases across all models. CNN and ResNet still maintain reasonable accuracy, but their scores drop compared to candlestick charts. Although the performance of ViT remains relatively stable across both image types in terms of accuracy, it still cannot match CNN and ResNet. The decline in performance when using bar-chart images suggests that candlestick charts contain richer visual patterns that are more informative for stock movement prediction, making them a better choice for deep learning-based forecasting.

Overall, the findings suggest that CNN and ResNet are well-suited for stock movement prediction when using candlestick chart images, while ViT struggles to extract meaningful patterns. In addition, the choice of stock image type significantly impacts model performance, with candlestick charts proving to be a more effective input representation than bar-chart images.

V. CONCLUSION AND LIMITATION

This study evaluated the effectiveness of the prompt-based method using the LLaMA model for predicting stock movements through two distinct approaches. The LLM performances are compared with CNN, ResNet and ViT models. The experimental results show that LLaMA has the potential to outperform traditional deep learning architectures (CNN, ResNet, and ViT) in predicting stock movement using

candlestick charts. In addition, a teacher-student approach with LLaMA and Qwen further improves forecasting performance across five stocks. In general, these findings suggest that LLMs can effectively capture complex stock price patterns from candlestick charts and surpass deep learning vision-based models in financial forecasting tasks.

However, this paper has some limitations that can be addressed and analyzed in future work. One such limitation is the class imbalance in the stock dataset, where the “Up” movements significantly outnumber the “Down” movements. This imbalance can negatively affect the model’s prediction performance, as it may lead to a bias toward predicting “Up” movements more frequently. Traditional resampling methods, while useful for handling imbalanced data, can disrupt the sequential structure of the time series, which is crucial for time series analysis.

In addition, this paper only uses five stocks for analysis. To represent a broader market, more stocks from different industries will be added for analysis in the future study.

For practical implementation, while LLMs have the potential to achieve strong performance, their computational and memory requirements are extremely high. Due to these limitations, this study utilizes only the “LLaMA-3.2-11B” model. Ideally, a larger model, such as “LLaMA-3.2-90B,” could deliver even better results.

CONFLICT OF INTEREST

The author declares no conflict of interest.

REFERENCES

- [1] G. Mu, N. Gao, Y. Wang, and L. Dai, “A stock price prediction model based on investor sentiment and optimized deep learning,” *IEEE Access*, pp. 353–367, 2023.
- [2] S. Wu, Y. Liu, Z. Zou, and T. H. Weng, “Si lstm: Stock price prediction based on multiple data sources and sentiment analysis,” *Connection Science*, vol. 34, pp. 44–62, Dec. 2022.
- [3] H. Kaeley, Y. Qiao, and N. Bagherzadeh, “Support for stock trend prediction using transformers and sentiment analysis,” 2023.
- [4] S. Wang, “A stock price prediction method based on bilstm and improved transformer,” *IEEE Access*, vol. 11, pp. 104 211–104, 2023.

- [5] H. Lee, J. H. Kim, and H. S. Jung, "Deep-learning-based stock market prediction incorporating esg sentiment and technical indicators," *Scientific Reports*, vol. 14, 10262, 2024.
- [6] A. Sharma, D. Bhuriya, and U. Singh, "Survey of stock market prediction using machine learning approach," in *Proc. 2017 International Conference of Electronics, Communication and Aerospace Technology (ICECA)*, 2017, vol. 2, pp. 506–509.
- [7] O. B. Sezer and A. M. Ozbayoglu, "Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach," *Applied Soft Computing*, vol. 70, pp. 525–538, 2018.
- [8] M. U. Gudelek, S. A. Boluk, and A. M. Ozbayoglu, "A deep learning based stock trading model with 2-d cnn trend detection," in *Proc. 2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2017, pp. 1–8.
- [9] C. C. Hung and Y.-J. Chen, "Dpp: Deep predictor for price movement from candlestick charts," *Plos One*, vol. 16, no. 6, e0252404, June 2021.
- [10] Q. Xie, W. Han, X. Zhang, Y. Lai, M. Peng, A. Lopez-Lira, and J. Huang. (2023). Pixiu: A large language model, instruction data and evaluation benchmark for finance. [Online]. Available: <https://arxiv.org/abs/2306.05443>
- [11] H. Yang, X. Y. Liu, and C. D. Wang. (2023). Fingpt: Open-source financial large language models. [Online]. Available: <https://arxiv.org/abs/2306.06031>
- [12] Z. Hu, J. Zhu, and K. Tse, "Stocks market prediction using support vector machine," in *Proc. 2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering*, vol. 2, pp. 115–118, 2013.
- [13] J. Zheng, D. Xin, Q. Cheng, M. Tian, and L. Yang, "The random forest model for analyzing and forecasting the us stock market under the background of smart finance," in *Proc. 3rd International Academic Conference on Blockchain, Information Technology and Smart Finance (ICBIS 2024)*, 2024, pp. 82–90.
- [14] K. Nakagawa and K. Yoshida, "Time-series gradient boosting tree for stock price prediction," *International Journal of Data Mining, Modelling and Management*, 2022.
- [15] A. John and T. Latha, "Stock market prediction based on deep hybrid RNN model and sentiment analysis," *Automatika*, vol. 64, no. 4, pp. 981–995, Oct. 2023.
- [16] B. A. Abdelfattah, S. M. Darwish, and S. M. Elkaffas, "Enhancing the prediction of stock market movement using neutrosophic-logic-based sentiment analysis," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 19, pp. 116–134, Jan. 2024
- [17] Q. Chen and H. Kawashima, "Stock price prediction using llm-based sentiment analysis," in *Proc. IEEE BigData 2024*, Washington DC, IEEE, 2024, pp. 4828–4835.
- [18] S. Bang and D. Ryu, "Forecasting stock prices using cnn-based visual representations of stock data," *Sungkyunkwan University Economic Research Institute, Tech. Rep.*, no. 23, 2023.
- [19] R. M. I. Kusuma, T. T. Ho, W. C. Kao, Y. Y. Ou, and K. L. Hua. (2019). Using deep learning neural networks and candlestick chart representation to predict stock market. [Online]. Available: <https://arxiv.org/abs/1903.12258>
- [20] Z. Li, S. Li, and X. Yan. (2023). Time series as images: Vision transformer for irregularly sampled time series. [Online]. Available: <https://arxiv.org/abs/2303.12799>
- [21] K. He, X. Zhang, S. Ren, and J. Sun. (2015). Deep residual learning for image recognition. [Online]. Available: <https://arxiv.org/abs/1512.03385>
- [22] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. [Online]. Available: <https://arxiv.org/abs/2010.11929>

Copyright © 2025 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).