Closing Price Prediction using Sentiment Analysis by LLM & SSA

Manan Baldwa[2], Vansh Jain[1], Parth Verma[1], Naman Kale[1], Umang Tripathi[1], Philip Treleaven[3]

[1]IIT Delhi, India [2]IIT Roorkee, India [3]University College London, UK

11 August 2024

Abstract

This project focuses on predicting the closing price of Tesla stock using sentiment analysis combined with a Long Short-Term Memory (LSTM) network and a novel Sparrow Search Algorithm (SSA) for model optimization. The sentiment analysis is performed by fine-tuning a Large Language Model (LLM) to extract sentiment scores from tweets related to Tesla, which are then used as a feature for the LSTM model. The project compares the performance of sentiment analysis methods and optimization techniques, demonstrating that fine-tuned LLMs and SSA provide significant improvements over basic methods.

Index Terms—Sentiment Analysis, LSTM, Large Language Model, Sparrow Search Algorithm, Stock Price Prediction

I Introduction

Predicting stock prices is a challenging task due to the dynamic and volatile nature of financial markets. Traditional methods often rely on historical price data and simple statistical models. However, integrating advanced sentiment analysis and modern optimization techniques can enhance prediction accuracy. This project explores the use of sentiment analysis via fine-tuned LLMs and SSA for optimizing LSTM models to predict Tesla's closing stock price.

II Data Collection

The dataset for this study includes Tesla stock price data and sentiment analysis derived from tweets. The stock price data was obtained from Yahoo Finance, covering 40 days, with data collected an hour before the market close. This dataset was divided into two segments of 20 days each: the first 20 days were used for training the models, while the remaining 20 days served as the test set.

Sentiment scores were extracted from tweets related to Tesla from Kaggle[1]. The tweets were filtered to ensure relevance and quality, focusing on those that contained substantial content about Tesla's corporate activities and market performance. This preprocessing step helped in obtaining more accurate sentiment insights for use in the prediction model.

III Methodology

A. Sentiment Analysis

VADER (Valence Aware Dictionary and Sentiment Reasoner) [2] is a sentiment analysis tool designed to analyze text and determine its sentiment polarity (positive, negative, or neutral). It is based on a predefined list of words and rules that assign sentiment scores to words and phrases. While VADER is useful for general sentiment analysis, it often struggles with the nuances and context-specific language in specialized datasets.

LLM (Large Language Model) [3] refers to advanced artificial intelligence models trained on vast amounts of text data to understand and generate human-like language. These models, when fine-tuned on specific datasets, can capture the subtleties and context of language more effectively than traditional methods.

Sentiment scores were extracted from tweets using a Large Language Model (LLM) that was fine-tuned specifically for Tesla-related content. The fine-tuning process employed QLoRA [4] with 4-bit precision on the first 20 days of data. This approach significantly improved performance, achieving an accuracy of 70% and a balanced accuracy of 67%, compared to an initial accuracy of just 46% before fine-tuning. This marks a substantial improvement over basic sentiment analysis methods like VADER. While VADER is a useful tool, it struggled with balanced accuracy and failed to capture the nuanced context of the tweets. Furthermore, a custom-made dictionary for sentiment analysis proved inadequate. Creating the dictionary was a manual, labor-intensive task that involved hand-picking relevant words from a dataset of 20,000 comments. Despite this effort, the dictionary fell short in capturing the complexity and context of the sentiments expressed, further highlighting the superiority of the LLM-based approach.

Sentiment scores were derived using a custom-designed function. Initially, an LLM generated logits, from which the predicted class was determined by taking the argmax. These logits were then passed through a softmax function to obtain probabilities. For bullish sentiment, the probability was multiplied by +1, while for bearish sentiment, it was multiplied by -1. Since multiple comments were associated with each date, the overall sentiment score was calculated as a weighted average. The weights were determined by the reciprocal of the class weight, ensuring a balanced contribution from each sentiment classification.

B. LSTM Model

Long Short-Term Memory (LSTM) [5] networks are a type of recurrent neural network (RNN)[5] designed to effectively capture long-term dependencies in sequential data. They achieve this by using memory cells and gates that regulate the flow of information, addressing the vanishing gradient problem seen in standard RNNs.

The choice of Long Short-Term Memory (LSTM) networks for stock price prediction is motivated by their superior ability to model sequential data and capture long-term dependencies. Stock prices exhibit temporal patterns where past prices, trends, and market conditions can significantly influence future values. LSTM models, with their unique architecture that includes memory cells and gated mechanisms, are particularly well-suited to handle such time-series data.

Compared to traditional models like ARIMA or even standard feedforward neural networks, LSTMs offer the advantage of retaining information over longer sequences, thereby capturing the temporal dependencies that are crucial for accurate stock price prediction. This ability to maintain context over time makes LSTMs a more robust choice for tasks involving financial data, where the underlying patterns can be complex and influenced by a variety of factors over extended periods.

In this study, an LSTM model was trained on stock market data, including features such as open, high, low, and volume. Hyperparameters were optimized using GridSearchCV [6], leading to an initial Mean Squared Error (MSE) of approximately 27.2

A column for Sentiment Scores was later added, and the LSTM was retrained with hyperparameters tuned again using Grid-

SearchCV. This resulted in an improved MSE of 22.15 with an overall average accuracy of about 25.175.

C. Sparrow Search Algorithm (SSA)

The Sparrow Search Algorithm (SSA) [7] is a bio-inspired optimization technique modeled after the foraging behavior of sparrows. It is designed to efficiently explore and exploit the search space, making it highly effective for optimizing complex models like LSTM networks. In this study, SSA is employed to optimize the hyperparameters of the LSTM model, which are crucial for improving its predictive performance.

SSA operates by simulating the behavior of two types of sparrows: producers and scroungers. Producers are responsible for searching for food (i.e., exploring the search space for potential solutions), while scroungers follow the producers to exploit their findings (i.e., fine-tune the solutions). This dual strategy allows SSA to balance exploration and exploitation, ensuring a comprehensive search across the hyperparameter space.

Initially, hyperparameter tuning was performed using GridSearchCV, a brute-force approach that resulted in a Mean Squared Error (MSE) of 22.15. While effective, GridSearchCV is computationally expensive and often less efficient at finding the optimal solution, especially in larger-scale problems.

In contrast, SSA significantly reduces computational time by focusing on the most promising regions of the search space, rather than exhaustively evaluating every possible combination of hyperparameters. This targeted approach not only accelerates the optimization process but also enhances the model's performance. The application of SSA led to a reduction in MSE from 22.15 to 17.2, demonstrating a 25% improvement over the grid search approach. This underscores SSA's superiority in providing quicker, more effective hyperparameter tuning for the LSTM model in stock price prediction.

IV Evaluation Metrics

The performance of the models is evaluated using three key metrics: Accuracy, Balanced Accuracy, and Mean Squared Error (MSE).

A. Accuracy

Accuracy [8] is a widely used metric in classification tasks, representing the proportion of correctly classified instances out of the total instances. It is mathematically defined as:

$$\begin{aligned}\text{Accuracy} &= \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \\ &= \frac{TP + TN}{TP + TN + FP + FN}\end{aligned}$$

where TP and TN are the true positives and true negatives, respectively, and FP and FN are the false positives and false negatives.

B. Balanced Accuracy

Balanced Accuracy [9] is particularly useful in datasets with imbalanced classes. It accounts for the accuracy of each class individually and then averages them, thereby providing a more nuanced evaluation. It is calculated as:

$$\text{Balanced Accuracy} = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right)$$

C. Mean Squared Error (MSE)

Mean Squared Error (MSE) [10] is utilized for regression tasks to measure the average squared difference between the predicted and actual values. For stock price prediction, MSE is expressed as:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

where n represents the number of data points, y_i represents the actual closing prices, and $\hat{y}_i$ represents the predicted closing prices.

In this research, Accuracy and Balanced Accuracy are applied to assess the performance of sentiment analysis from the large language model (LLM). Mean Squared Error (MSE) is used to evaluate the predictive performance of the stock price models.

V Results

The results of the various sentiment analysis and prediction approaches are summarized below:

Model	Balanced Accuracy	Accuracy
Sentiment VADER	51%	76%
LLM without Fine-Tuning	46.23%	40.4%
LLM with Fine-Tuning	67.3%	69%

Table 1: Balanced Accuracy and Accuracy of different sentiment analysis models.

Model	MSE
LSTM	27.2
LSTM + Sentiment Score	22.15
LSTM + Sentiment Score + SSA	**17.2**

Table 2: Mean Squared Error (MSE) of different LSTM-based models.

VI Observations

The results highlight that the fine-tuned Large Language Model (LLM) provides a 20% improvement in accuracy compared to the basic VADER sentiment analysis, significantly enhancing prediction performance. Additionally, the integration of the Sparrow Search Algorithm (SSA) further optimizes the LSTM model, achieving a 25% reduction in Mean Squared Error (MSE) compared to the grid search approach. This underscores the effectiveness of both advanced sentiment analysis techniques and efficient hyperparameter optimization in delivering more accurate and robust stock price predictions.

VII Conclusion

This project highlights the effectiveness of combining advanced sentiment analysis with state-of-the-art optimization techniques to enhance stock price prediction models. The use of a fine-tuned Large Language Model (LLM) for sentiment analysis significantly improves the model's ability to capture and interpret complex financial sentiments. Concurrently, applying the Sparrow Search Algorithm (SSA) for hyperparameter optimization of Long Short-Term Memory (LSTM) networks leads to notable improvements in model performance.

These advancements illustrate the substantial benefits of integrating sophisticated analytical tools and optimization strategies for achieving more accurate and reliable financial predictions. The success of these methods demonstrates their potential for improving forecasting accuracy and opens up opportunities for further exploration and application in financial analytics.

VIII References

[1] Frank Cao Yun "Stocktwits 2020-2022 Raw". Available: `https://www.kaggle.com/datasets/frankcaoyun/stocktwits-2020-2022-raw`

[2] Shihab Elbagir and Jing Yang "Twitter Sentiment Analysis Using Natural Language Toolkit and

VADER Sentiment ” Available: https://www.iaeng.org/publication/IMECS2019/IMECS2019_pp12-16.pdf

[3] Humza Naveed, Asad Ullah Khan, Shi Qiu, Muhammad Saqib, Saeed Anwar, Muhammad Usman, Naveed Akhtar, Nick Barnes, Ajmal Mian ”A Comprehensive Overview of Large Language Models” Available: https://arxiv.org/pdf/2307.06435

[4] Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, Luke Zettlemoyer ”QLoRA: Efficient Finetuning of Quantized LLMs” Available: https://arxiv.org/abs/2305.14314

[5] Greg Van Houdt, Carlos Mosquera , Gonzalo Nápoles ”A Review on the Long Short-Term Memory Model” Available: https://www.researchgate.net/publication/340493274_A_Review_on_the_Long_Short-Term_Memory_Model

[6] Petro Liashchynskyi, Pavlo Liashchynskyi ”Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS” Available : https://arxiv.org/abs/1912.06059

[7] Jiankai Xue and Bo Shen ”A novel swarm intelligence optimization approach: sparrow search algorithm” Available: https://www.tandfonline.com/doi/full/10.1080/21642583.2019.1708830#abstract

[8] Antonio Menditto, Marina Patriarca, Bertil Magnusson ”Understanding the meaning of accuracy, trueness and precision” Available: https://www.researchgate.net/publication/226202473_Understanding_the_meaning_of_accuracy_trueness_and_precision

[9] Kay Henning Brodersen, Cheng Soon Ong, Klaas Enno Stephan, Joachim M. Buhmann ”The Balanced Accuracy and Its Posterior Distribution ” Available: https://ieeexplore.ieee.org/document/5597285

[10] Timothy O. Hodson, Thomas M. Over, Philip K. Chan, Sydney Foks ”Mean Squared Error, Deconstructed” Available: https://www.researchgate.net/publication/356500557_Mean_Squared_Error_Deconstructed

[11] Guangyu Mu, Nan Gao, Yuhan Wang, and Li Dai ”A Stock Price Prediction Model Based on Investor Sentiment and Optimized Deep Learning” Available: https://ieeexplore.ieee.org/abstract/document/10130578

[12] Narayana Darapaneni, Anwesh Reddy Paduri, Himank Sharma, Milind Manjrekar, Nutan Hindlekar, Pranali Bhagat, Usha Aiyer, and Yogesh Agarwal ”Stock Price Prediction using Sentiment Analysis and Deep Learning for Indian Markets” Available : https://arxiv.org/pdf/2204.05783

[13] Adriano Koshiyama, Nick Fioozye and Philip Treleaven ”Algorithms in Future Capital Markets” Available : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3527511