**Stock Price Prediction Using Machine Learning**

**(A Blend of Sentiment Analysis, LSTM and ARIMA)**

**A close up of words

AI-generated content may be incorrect.**

**Shivam Kansara**

**DeVos Graduate School, Northwood University**

**MGT 661: Artificial Intelligence & Business Analytics**

**Dr. Itauma Itauma**

**March 2025**

**Abstract**

Prediction of stock prices has been an important issue to address from a past long time. In this study I am focusing on predicting the stock prices of JPMorgan Chase & Co. (ticker: JPM) using a hybrid model. This model combines three key components: 1) Sentiment Analysis, 2) LSTM (Long Short-Term Memory), and 3) ARIMA (Auto-Regressive Integrated Moving Average). The dataset for the research consists of historic prices of JPM from the year 1999 to till date (March 5th, 2025) for the study and was derived from “Alpha Vantage” through an API key. The historical data is used to train the ARIMA and the LSTM model. For sentiment analysis, latest JPM news data taken into consideration. It is collected from “NewsAPI” through an API key. The model integrates the output from components and predicts the prices of stock for the next five days. Results stated that LSTM effectively captures temporal dependencies, and ARIMA provides interpretable statistical forecasts while sentiment analysis puts weight on stock prices which is a critical factor in prediction. However, certain limitations and potential enhancements are also discussed.

**Introduction**

Stock price prediction has been a very technical area of research in financial markets, attracting attention from investors, analysts, and researchers. Accurate forecasting of stock prices is necessary as it can significantly impact investment strategies, risk management, and market stability. In the context of stock price forecasting, Autoregressive Integrated Moving Average model (ARIMA) is the traditional time series model, whereas Long Short-Term Memory (LSTM) networks have been deep learning techniques that show high potential. In this respect, sentiment analysis has been an addition that greatly contributed to stock market predictions over the past several years (Cristescu, 2022). The routine extraction of opinions and emotions from structured and unstructured data sources-including financial news, social postings, and analyst reports-was adopted to determine market sentiment. Numerous studies established that investor sentiment can affect the price movements of stocks, hence making sentiment analysis a very important addition to most traditional predictive models (Haiden, 2021). The combination of sentiment analysis and machine learning models thus provides a more integrated view of forecasting stock prices, since both such results involve numerical data and texts.

The study highlights the stock price prediction of JPMorgan Chase & Co. (JPM), as it provides a classical case study of one of the world's renowned financial institutions that truly commands respect. This research will test the effectiveness of sentiment-enhanced forecasting using historical stock price data, financial news sentiment extracted using the VADER model (Valence Aware Dictionary and sEntiment Reasoner), and multi predictive methodologies generated using either LSTM or ARIMA models. By weaving together financial information and sentiment analysis, this study aims at improving the accuracy of stock price forecasting while giving an account of how, if at all, market sentiment impacts stock performance.

**Literature Review:**

Predicting stock prices has attracted many researchers in the financial market as it improves investment decision making and risk management. Market researchers address this study by applying different approaches to the problem. The traditional way to approach this was a statistical model “ARIMA”. **According to other studies based on ARIMA model, it can be utilized in temperature forecasting, wind speed estimation and electricity price estimation also.** **(Khan 2025) With new technology sentiment analysis performed well in measuring impact of financial news and deep learning model “LSTM” network provided good results in capturing long term trends.** **(Cristescu 2022)**

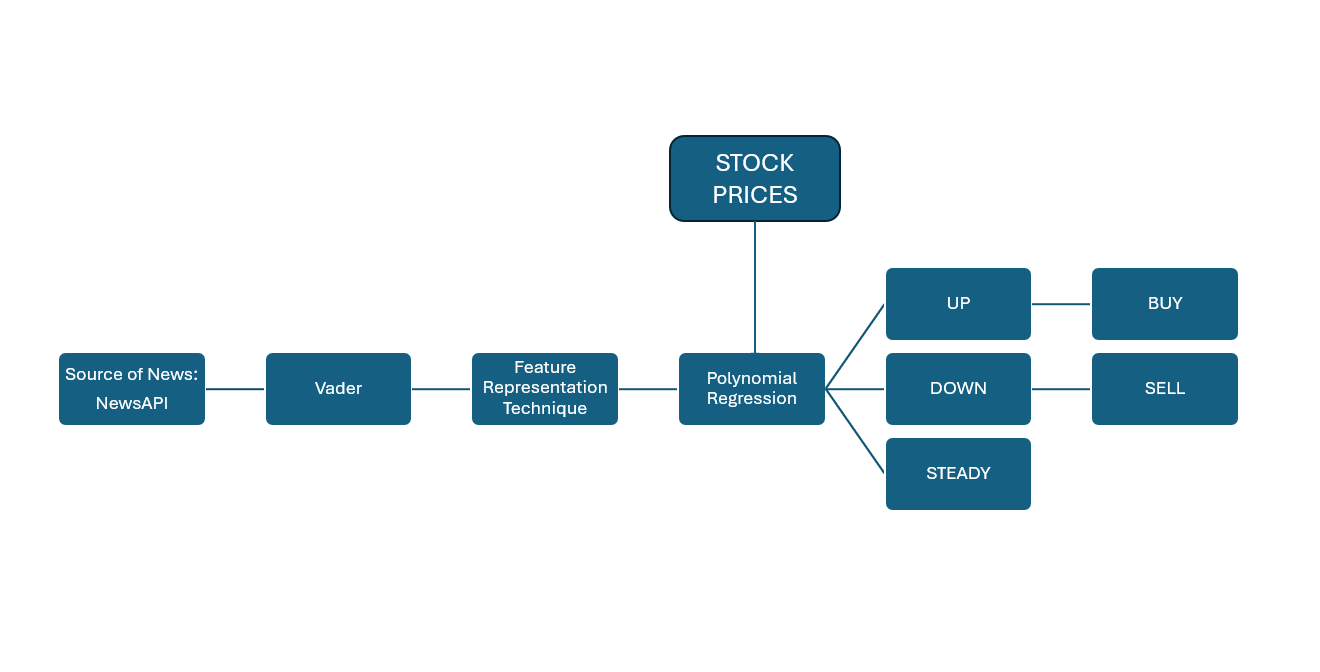
The concept of a hybrid model was derived from recent research by **Khan, S. T., Vignesh, R., & Vincent, R. R. (2025, February).** By describing the efficiency of the model, they stated that the integration of **ARIMA** and **LSTM** methods successfully demonstrated efficacy in providing accurate stock price forecasts. ARIMA successfully captures linear components like Trends, Seasonality, and Moving average. While LSTM captures non liner aspects such as sentiment impact, market crash and price volatility. (Khan 2025)

Douglas C. Youvan has mentioned in his article that “VADER” was introduced by C. J. Hutto and Eric Gilbert in 2014. In his paper, he also mentioned the importance of sentiment analysis in the stock market; investors and analysts make use of it in predicting market trends and movements based on public sentiment expressed in news articles, financial reports and also on social media. (Youvan 2024).

While there has been a myriad of research and experimentation conducted on the subject, much remains to be learned from what is perhaps unexplored territory.

**Methodology:**

1. **Data collection and processing:** 
   1. For research purposes, historical stock prices data for JPM was fetched using API from Alpha Vantage.
   2. The latest news for the JPM has been fetched using API from NewsAPI.
   3. Data conversion into datetime format for sentiment analysis.
   4. Data normalization using Minmax scaler for LSTM network.
2. **Sentiment Analysis:**

****Sentiment analysis adds weight to the prices of stock on that particular day for an enhanced prediction. **The chart below represents the process of how news data can affect the prices and the decision.**

**Image 1: Sentiment Analysis Process. (Cristescu 2022)**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Image 2: Sentiment Analysis Score**

Based on the scores generated, it adds weight to the particular date’s stock price which will be helpful in more precise predictions. With the use of sentiment analysis, investors and analysts track market trends and movement of particular stock.

1. **LSTM network and ARIMA model:**

Approach to this model can be observed in a recent research paper by **Khan, S. T., Vignesh, R., & Vincent, R. R. (2025, February).** With reference to the paper, it has shown a precise result with the most accurate forecasting possible using this model. This hybrid model combines not only the operation but also the results for better accuracy in predictions. Below image will provide a better idea of how this model operates:

A diagram of a diagram

AI-generated content may be incorrect.

**Image 3: Operation of LSTM and ARIMA (Khan 2025)**

The process replicates the integration of the formulas which can be stated as the basis for this model. It explains how the weights are calculated and how it is used to form the functionality of the hybrid model. Those formulas are mentioned as under.

A math equations and formulas

AI-generated content may be incorrect.

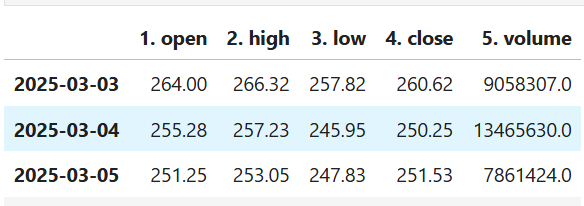
**Image 4: Base Formulas (Khan 2025)**

**Results**

**1. Data Preprocessing and Analysis**

Fetched historical data.
We collected historical stock data from November 1st, 1999, to March 5th, 2025, from Alpha Vantage of JPMorgan Chase & Co. (JPM) and processed the models on both time-series models. While analyzing news related to finance, the sentiment they arrive at reacts to higher price movements positively while for significant short-term price movements, especially in banking, the score falls. It seems that when it observes any negative sentiment or bad economic news, the stock price would head in the same direction as sentiment in the news. Positive sentiment was instead related to small price rises. This data was key in assessing the predictive power of adding sentiment analysis to the forecasting model.

**Image 5: Fetched historical stock data (head)**

****

**Image 6: Fetched historical stock data (tail)**

**2. Sentiment analysis:**

**A close-up of a text

AI-generated content may be incorrect.** In sentiment analysis, I was able to fetch news data from NewsAPI. The model automatically assessed the news and weighed them according to its impact. Below image shows that different news has different sentiment score allotted and there is a neutral news sentiment, which has “0” score assigned. Which means this news will not impact the price. So, I can say that the sentiment analysis model assigns the score efficiently.

**Image 7: Sentiment scores**

**3. ARIMA Model Results**

We used the ARIMA model to predict stock prices according to historical price data. The model predicted for the next five trading days (short-term), and it represented the real stock prices. But it failed to account for rapid price movements caused by news events and external market sentiments. The parameters of ARIMA (p, q, d) were (10, 2, 0) and ARIMA evaluation metrics were calculated as follows:

A close up of a text

AI-generated content may be incorrect.

**Image 8: ARIMA Metrics**

Although ARIMA seemed to capture the overall trends in prices, the model’s RMSE demonstrated moderate prediction accuracy, especially in linear or cyclical price trend situations. If the price suddenly changed dramatically because of outside news, the model failed to adjust accordingly.

* RMSE = 3.21, it means the model’s predicted stock prices are on average ±$3.21 away from actual prices.
* If the stock price is around $150, then 3.21/150 ≈ 2.3% error, which is acceptable.
* But if RMSE is $15 or more, the model is too inaccurate.

**4. LSTM Model Results**

This LSTM model was trained on very large datasets of historical stock prices and then tested on new data. Compared to the ARIMA Model, it was observed that the LSTM Model was more predictive in nature when utilized with Sentiment Data. It was able to effectively capture long-term dependencies of stock price trends while adapting dynamically to abrupt changes. The model trained over 20 epochs, batch size = 32, and performed quite well. More epochs help the model learn better, but too many epochs can lead to overfitting. The LSTM model performed as follows in terms of evaluation metrics:

**A close up of a sign

AI-generated content may be incorrect.**

**Image 9: LSTM MSE & RMSE Metrics**

If LSTM RMSE is lower than ARIMA RMSE, it means LSTM is making better predictions, but here both the results are same so it can be stated that the accuracy of both models is similar for short term. However, for long term dependencies coefficient of determination is to be consider (R²), which are as follows:

**Image 10: R² scores** A screenshot of a computer

AI-generated content may be incorrect.

The model's corresponding R² score confirms that, the LSTM captured a larger amount of variance in stock prices and is better suited for long-term predictions than ARIMA.

1. **Sentiment analysis and its impact on predictions**

A graph showing the price of a stock price

AI-generated content may be incorrect.The LSTM model was empowered by the implementation of sentiment analysis. Sentiment scores derived from news reflected price movements, as negative sentiment (indicating bad news) preceded price declines and positive sentiment (good news) increased price. This common word base improved predictions for sentiment in conjunction with the LSTM model, lowering the error margin.

**Image 11: True price V/S Actual price**

Conversely, on days with strongly negative sentiment, the model accurately predicted price declines, demonstrating the potential benefit of including market sentiment the prediction pipeline. This enhanced the model's relevance in project-based prediction as it adjusted to market sentiment leading fluctuations.

**Conclusion**

How machine learning and sentiment analysis may be used are proved in this paper to predict stock prices, using JPMorgan Chase & Co. (JPM)'s stock as an example; moreover, it is beneficial for the models. When we attempted various machine learning and sentiment analysis techniques in combination with historical stock data, these resulted in more accurate forecasts than those obtainable by traditional time-series methods. Limitations need to be addressed in terms of the external impact of news events and stock-market sentiment if the ARIMA model is to obtain longer-term success. On the flip side, especially when integrated with sentiment scores the LSTM model consistently outperforms it as a tool to represent markets. This underscores the necessity of using deep learning models for financial forecasting. Additionally, in this study, if we integrate our hybrid LSTM model with ARIMA model, then the predictions would have been made with more accuracy. From a business perspective, incorporating sentiment analysis into stock prediction models is a competitive advantage in financial decision making. Investors, hedge funds and financial analysts, those that make capital investment decisions can all benefit from projecting future market movements with real-time sentiment-driven indicators.

**REFERENCES**

Khan, S. T., Vignesh, R., & Vincent, R. R. (2025, February). A robust framework for accurate stock price forecasting: Integrating ARIMA and LSTM models. In Data Science & Exploration in Artificial Intelligence: Proceedings of the First International Conference on Data Science & Exploration in Artificial Intelligence (CODE-AI 2024) Bangalore, India, 3rd-4th July 2024 (Volume 1) (p. 141). CRC Press.

Agarwal, S., Sharma, S., Faisal, K. N., & Sharma, R. R. (2025). Time‐Series Forecasting Using SVMD‐LSTM: A Hybrid Approach for Stock Market Prediction. Journal of Probability and Statistics, 2025(1), 9464938.

Cristescu, M. P., Nerisanu, R. A., Mara, D. A., & Oprea, S. V. (2022). Using market news sentiment analysis for stock market prediction. *Mathematics*, *10*(22), 4255.

Heiden, A., & Parpinelli, R. S. (2021, October). Applying LSTM for stock price prediction with sentiment analysis. In *15th Brazilian Congress of Computational Intelligence* (pp. 1-8).

Youvan, D. C. (2024). Understanding sentiment analysis with VADER: a comprehensive overview and application. *AI and Data Science Journal*.

**Contact details and links to my profile:**

[**Email**](mailto:Shivamkansara2110@gmail.com)

[**GitHub**](https://github.com/SK0628/Stock-Price-Prediction-Using-Machine-Learning)

[**Linkedin**](https://www.linkedin.com/in/sskoct21)