

Correlating Stock Market News Sentiment with NVIDIA Share Price Movements

William Giles, William Mitchell, Chai Zhang



UVA

SCHOOL *of* DATA SCIENCE



1. **Motivation & Goal**
2. **Data Acquisition**
3. **Analysis Plan**
4. **Bias & Uncertainty**
5. **Results**
6. **Next Steps**

MOTIVATION

This project investigates how financial news sentiment influences stock market behavior, focusing specifically on NVIDIA stock. Due to the recent surge in news coverage and significant price fluctuations, NVIDIA presents an ideal subject for study.

Hypothesis: There has been a correlation between the sentiment of news articles regarding NVIDIA and the company's opening share price the following day within the month of January.

This allows for us to analyze and draw comparisons between news sentiment scores and opening share prices, which could be transformed and used in a variety of financial applications.

GOALS

Research Question: Can the direction of NVIDIA stock price trends be accurately predicted using sentiment analysis models based on financial news sentiment?

Modeling Approach: Using Alpaca API, we will gather NVIDIA's share price data and sentiment analysis data from news articles regarding NVIDIA. Using OpenAI, a sentiment analysis will be performed to determine negative, neutral, or positive sentiment. The opening share price and change in share price per day will be plotted. These share prices and sentiment scores will be compared to determine correlation.

Our goal using this Modeling Approach is to determine whether sentiment analysis models can accurately predict the direction of NVIDIA's stock price based on the sentiment expressed in financial news.

DATA ACQUISITION

Using the Alpaca API, we collected financial news data and price data for the NVIDIA stock.

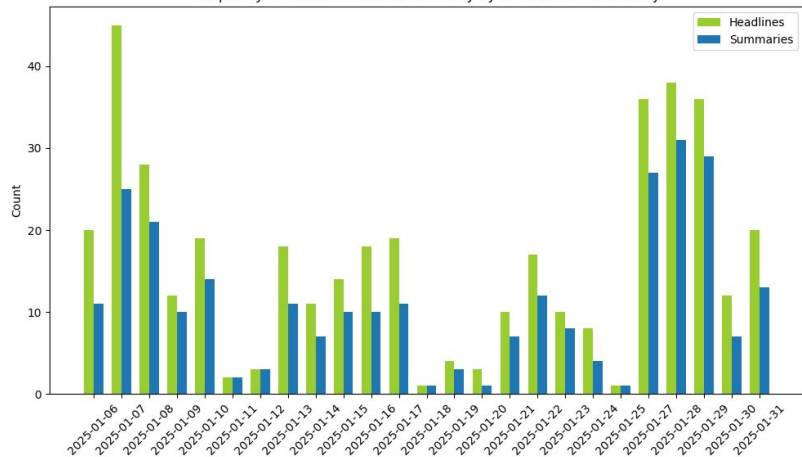
News Data:

- Retrieved daily news articles for NVIDIA from January 6 to January 31, 2025 using the Alpaca News API
- Extracted key fields: publication date, headline and summary
- Aggregated the cleaned data into files for further analysis

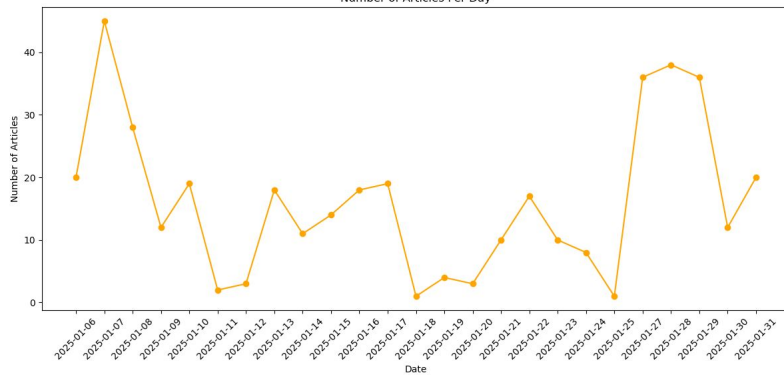
Stock Price Data:

- Retrieved daily price bar data for NVIDIA over the same period using Alpaca API
- Extracted the date and the opening price for each day

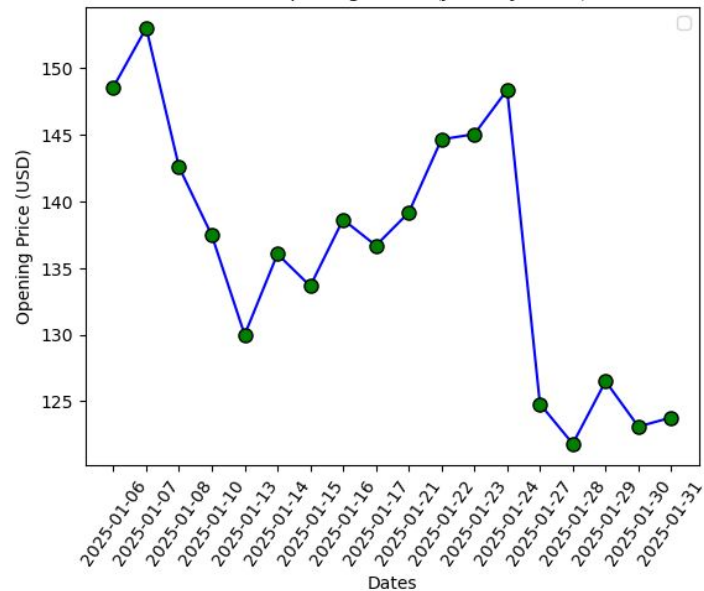
Frequency of Articles Published Each Day by Headline and Summary

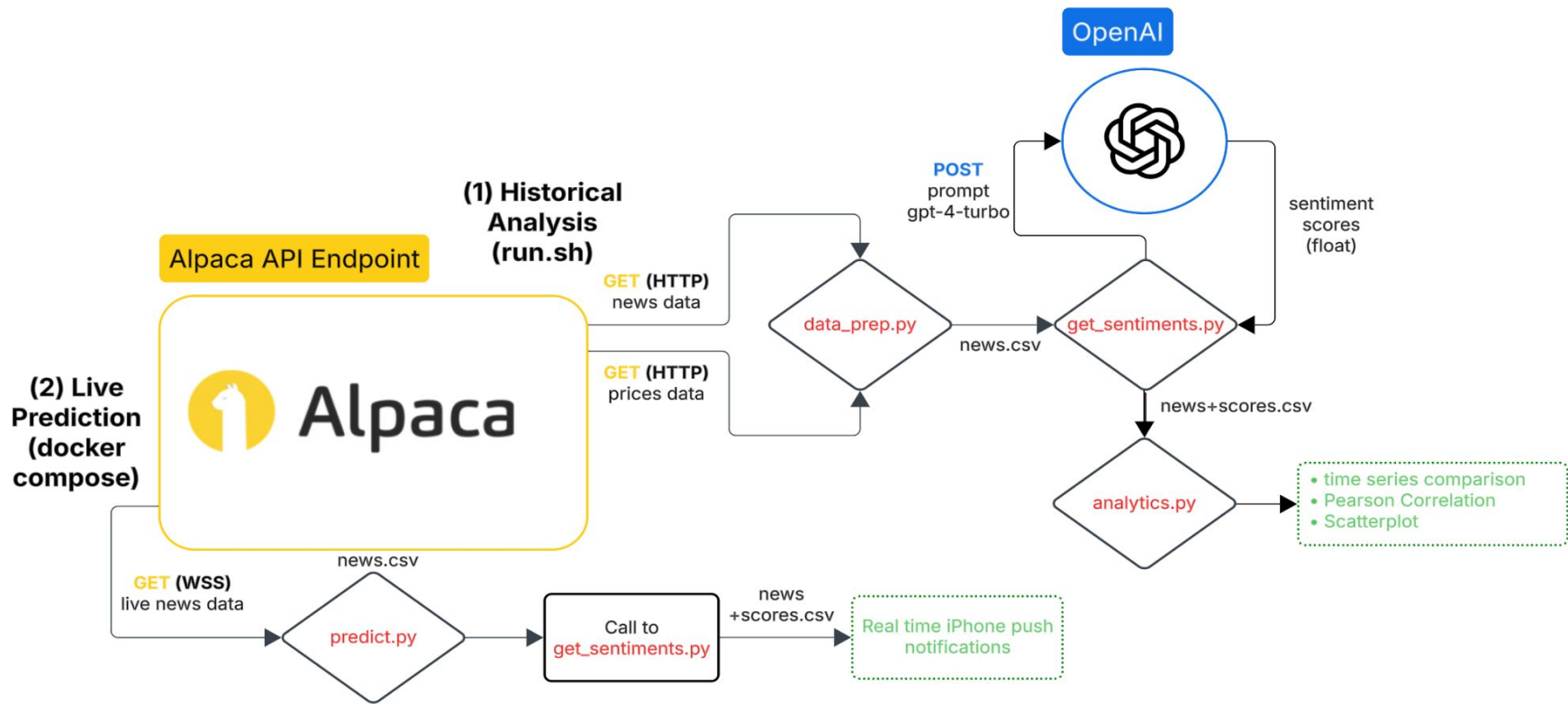


Number of Articles Per Day

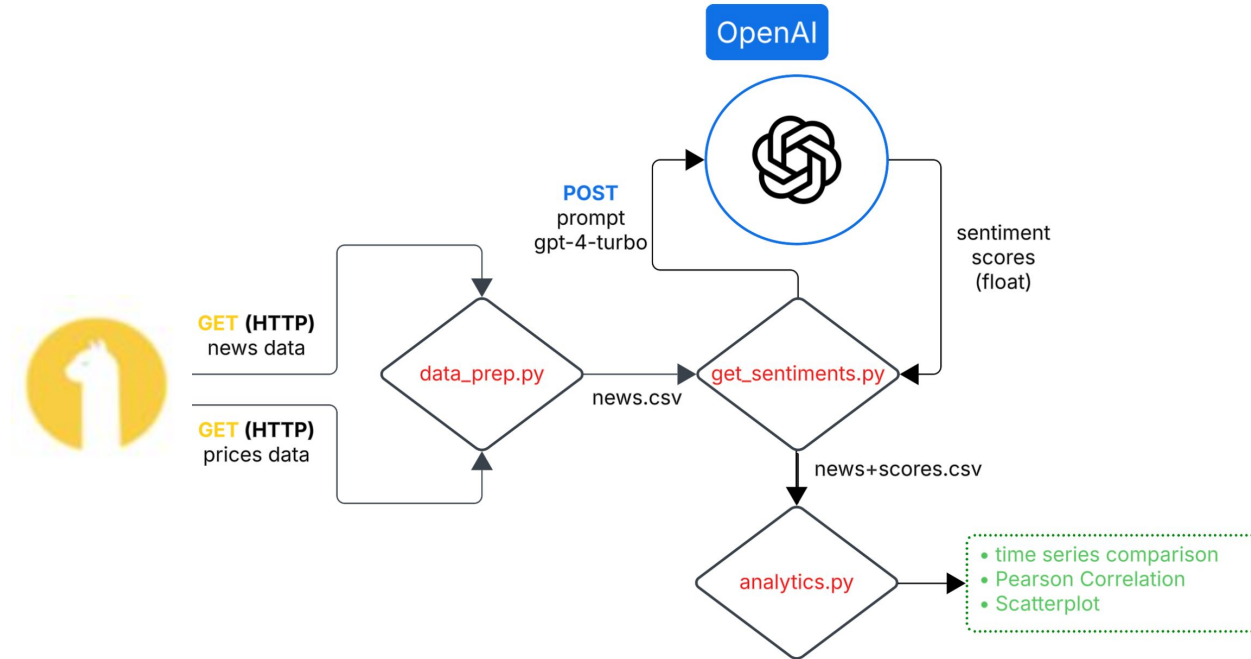


NVIDIA Opening Prices (January 2025)

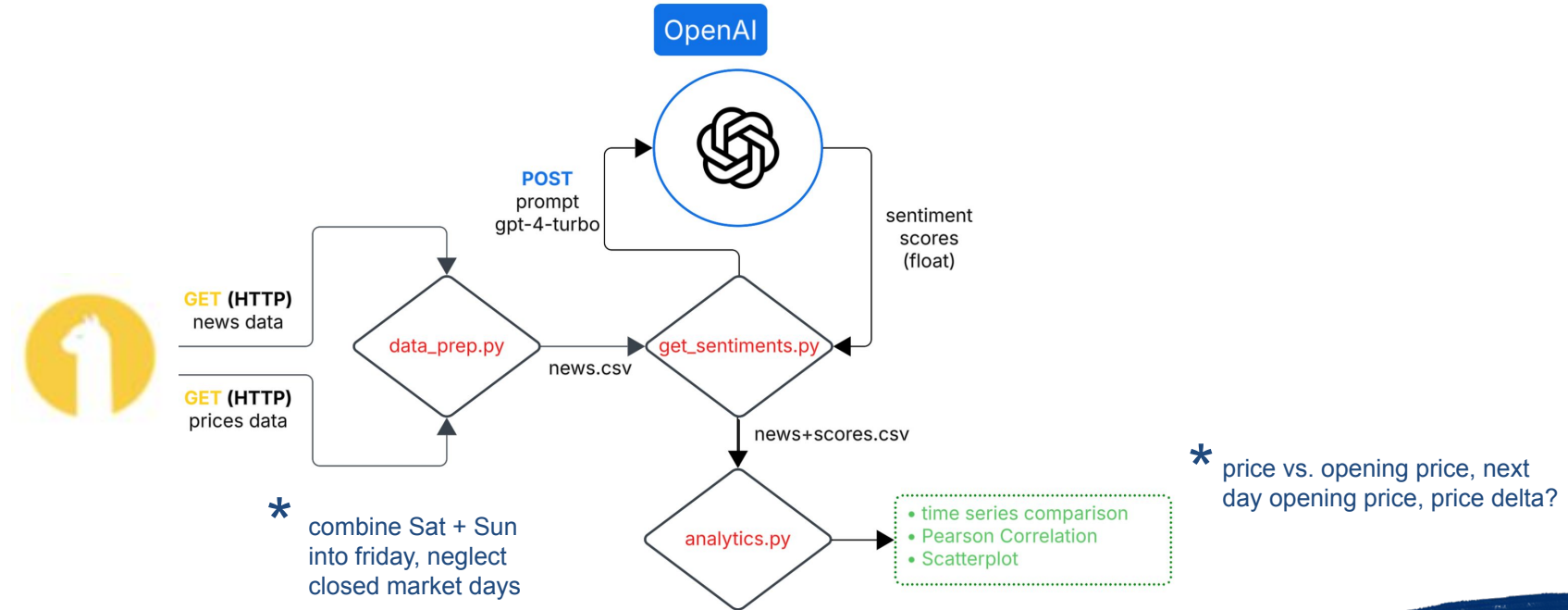




PIPELINE I. HISTORICAL ANALYSIS (run.sh)



PIPELINE I. HISTORICAL ANALYSIS (run.sh)



PIPELINE II. LIVE PREDICTION (docker compose)



GET (WSS)
live news data

news.csv

predict.py

Call to
get_sentiments.py

news
+scores.csv

Real time iPhone push
notifications

BIAS & UNCERTAINTY

Data Bias:

- NVIDIA was the only stock analyzed through this study, and through this only the opening share prices. and change per day were analyzed.
- The month of January was the only time frame used for the study.

Model Bias:

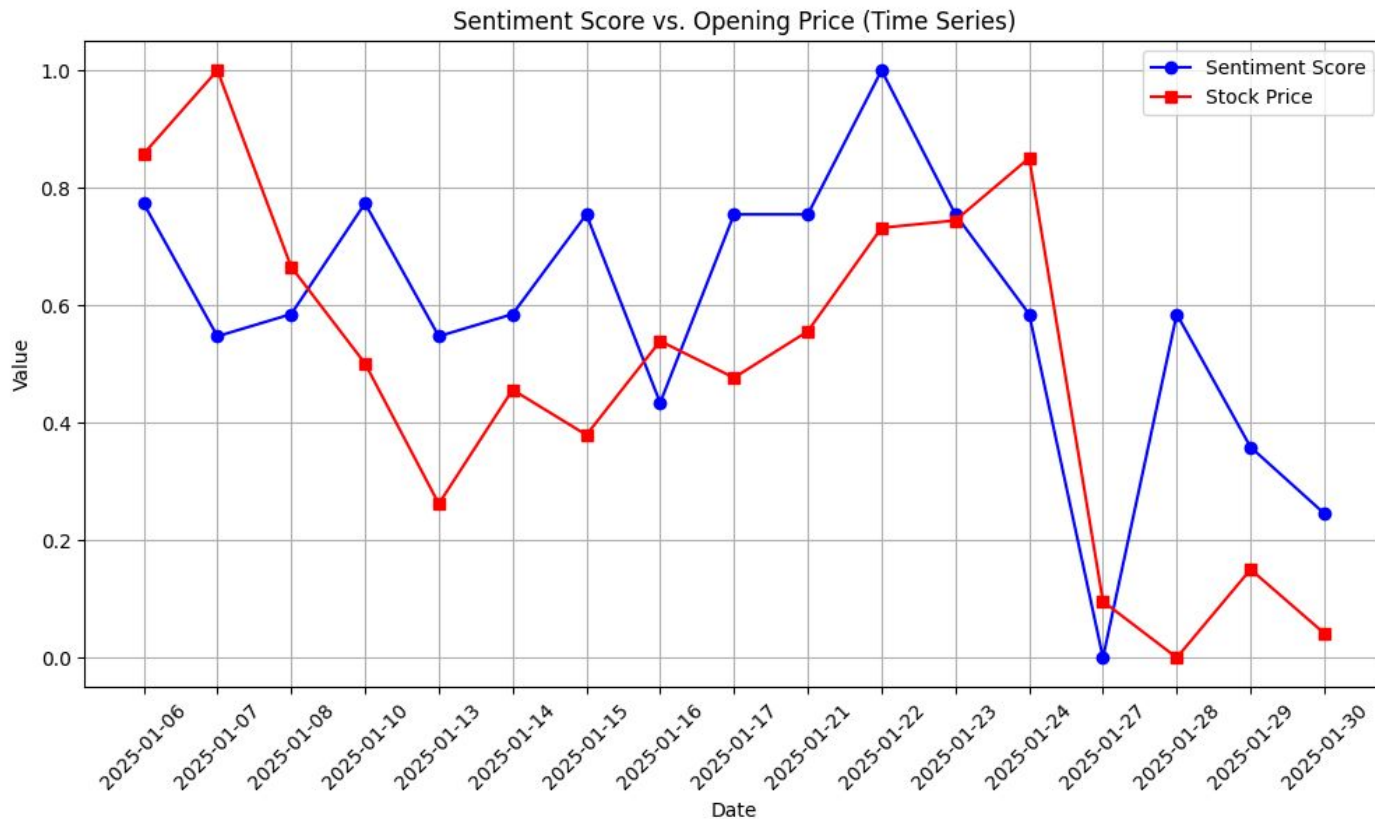
- Only OpenAI was used, and there are concerns about it analyzing more than just the news. There is potential that it has access to prices to sway the sentiment score that it gives.

The uncertainty in the values that we report are due to these biases, with potential for the data values to be misconstrued. With correction of these biases through expansion of the data set and refining our model, we hope to eliminate them and have stronger value certainty.

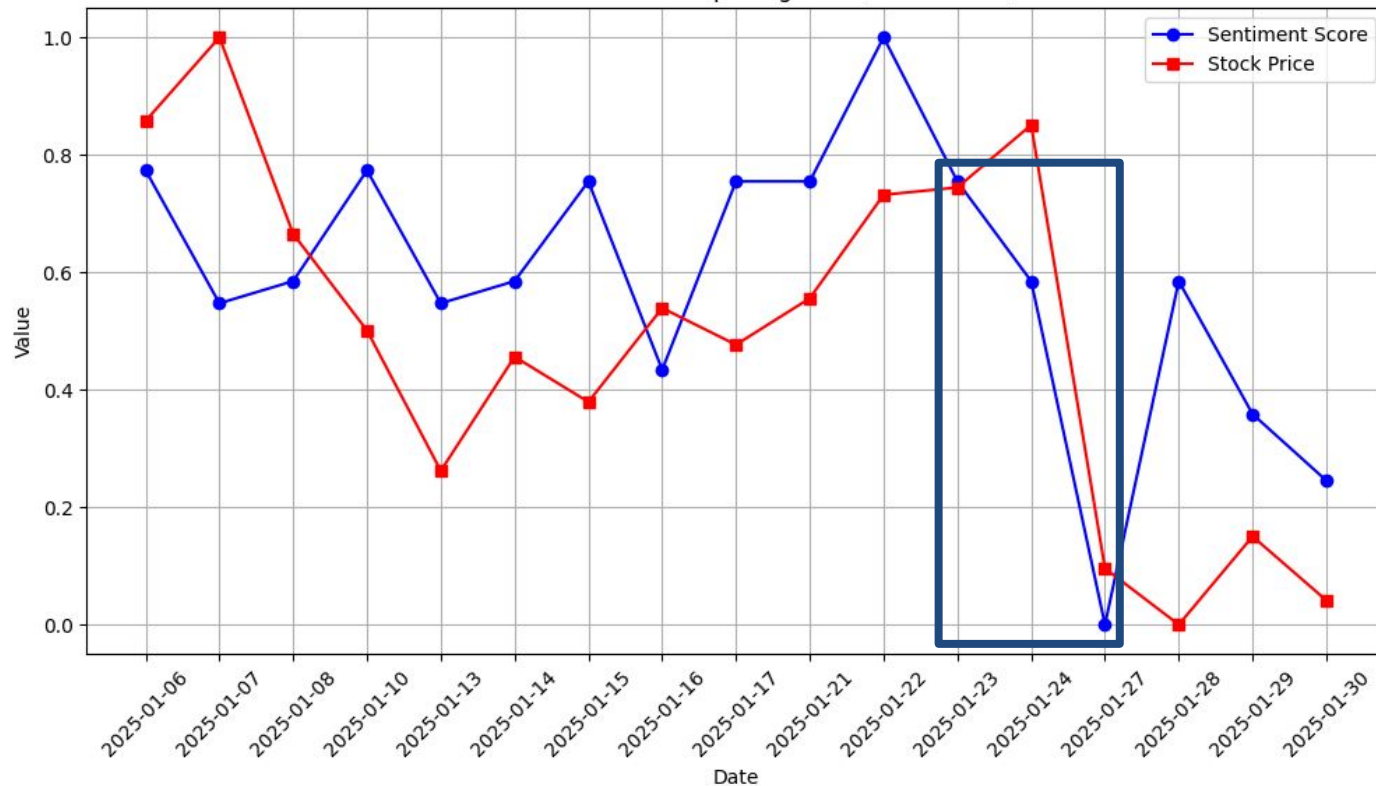
RESULTS

The next day's opening price of NVIDIA exhibits a strong correlation ($r = 0.71$) with sentiment in financial news, suggesting that news sentiment could serve as a valuable indicator for investor decision-making regarding NVIDIA.

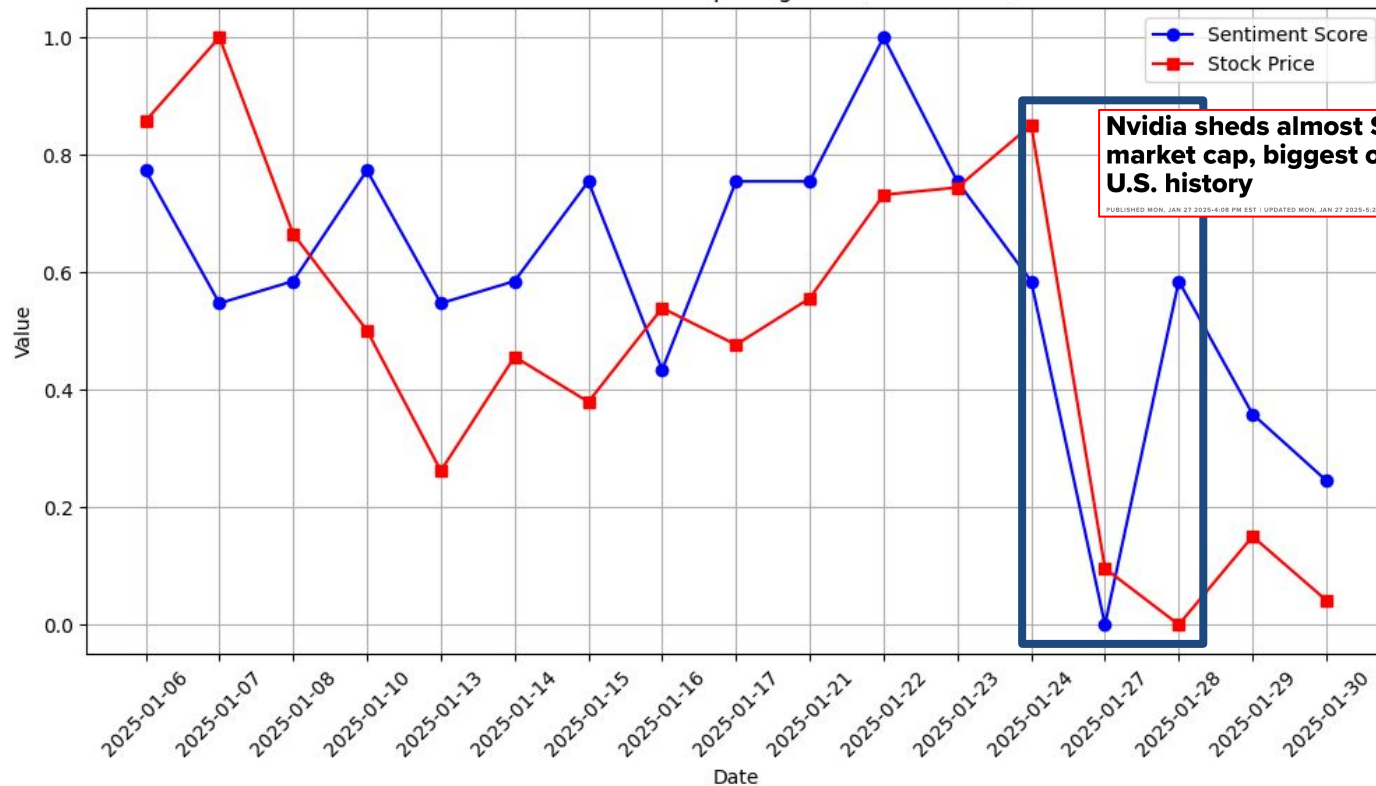
Therefore, our hypothesis is supported by our results, as there is indeed a strong correlation in the month of January.



Sentiment Score vs. Opening Price (Time Series)

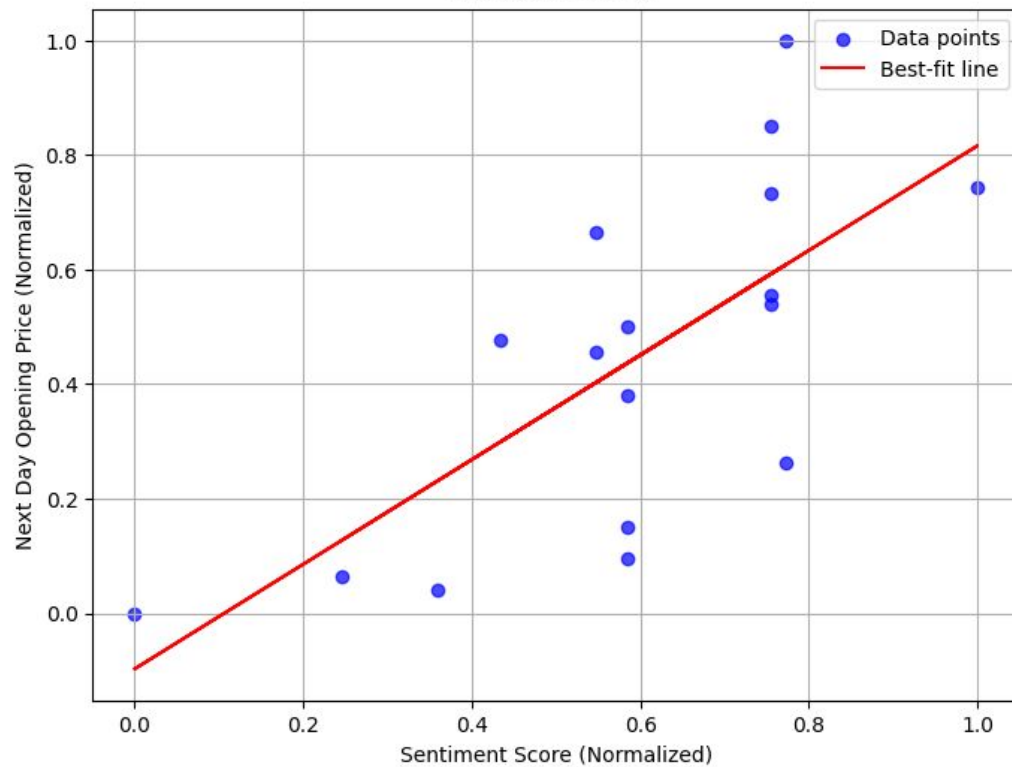


Sentiment Score vs. Opening Price (Time Series)



Scatter Plot of Sentiment Score vs. Next Day Opening Price

Correlation: 0.71



NEXT STEPS

Data Acquisition

- Collect data over a longer period of time (> 1 month)
- Collect data for various symbols (AMZN, TSLA, etc...)
- Are prices of certain asset classes and market segments more or less correlated with their news data?

Modeling Approach

- Further fine-tune gpt-4-turbo
- Experiment with other models: openai, other open source, fine-tuned financial sentiment models to validate results

Live Deployment

- Bolster architecture supporting live inference
- How can we use this correlation to best guide investment decisions?

REFERENCES

[1] A. Sharma, "Fine-tuning Large Language Models," *Analytics Vidhya*, Aug. 2023. [Online]. Available: <https://www.analyticsvidhya.com/blog/2023/08/fine-tuning-large-language-models/>. [Accessed: Jan. 30, 2025].

[2] Alpaca, "News API Reference," *Alpaca Documentation*, [Online]. Available: <https://docs.alpaca.markets/reference/news-3>. [Accessed: Feb. 16, 2025].

[3] L. Bartolini, L. Goldberg, and A. Sacarny, "How Economic News Moves Markets.," *Current Issues in Economics and Finance*, vol. 14, no. 6, August 2008. [Online]. Available: https://www.newyorkfed.org/medialibrary/media/research/current_issues/ci14-6.pdf [Accessed: Jan. 30, 2025].

[4] "The data appendix," The Data Appendix | Project TIER | Teaching Integrity in Empirical Research, <https://www.projecttier.org/tier-protocol/protocol-4-0/root/data/analysisdata/data-appendixfile/> [Accessed: Feb. 16, 2025].

Github: <https://github.com/wdgstl/MarketSentimentAnalysis.git>

Sentiment Score vs. Opening Price (Time Series)

