MDA 720 CAPSTONE

Social Media's Impact on the stock market

Andre Kleppe

LIU Post / Gurperet Singh

Table of Contents

Introduction and background

Social Sentiment Influence: Measuring Stock Movements through Social Media Analytics.

In the age of social media, outlets like X, Reddit, and Telegram have evolved from casual networking hubs to significant drivers of financial market "advice" and discussion. Individual users and "influencers" generate massive visibility and attention around stocks and financial products in minutes, creating surges in investor interest and triggering price volatility. Events such as the GameStop (GME) short squeeze and viral tweets by Elon Musk and Donald Trump exemplify the role of social media sentiment in shaping short-term investor behavior and disrupting traditional market signals.

This project aims to investigate the impact of social media sentiment on stock performance. It explores not only how public opinion, expressed in posts, hashtags, and trending discussions, correlates with short-term price changes but also how companies might use social media to their advantage. The goal is to determine whether organizations can use media and the public to optimize their market strategy and potentially drive outcomes that are financially favorable.

Through web scraping, natural language processing (NLP), and market data analysis, this study extracts sentiment signals from posts on Reddit through Praw. Using models like VADER sentiment analysis, where sentiment scores are computed and aligned with corresponding stock price data from Yahoo Finance and Alpha Vantage. The resulting dataset allows us to assess real-time sentiment shifts, their intensity, and any lagged or immediate impact on a stock price.

Not only will this project be about descriptive analysis, it will also consider a forward-looking application: how companies could shape narratives to align with their goals. This extends the scope from analysis and statistics into actionable strategy, empowering investor relations teams, marketers, and executives with tools to influence the public.

Research Objective

The main objective for this project is to analyze how social media sentiment correlates with short-term stock price movements and to explore how organizations can strategically adapt their marketing efforts to better reach their audience. As social media is now used to checking political updates, to what the aunt in Australia did last week. This is a golden opportunity we can explore, which hopefully will positively benefit us

By analyzing sentiment from user-generated posts and comparing it to corresponding stock movements in the market, this study aims to deliver data-driven insights that bridge marketing strategy with financial analytics.

Research Questions

Correlation:

To what extent does social media sentiment, as measured through NLP techniques, correlate with short-term stock price fluctuations?

Strategic Marketing:

Can companies use real-time sentiment trends to adjust their messaging or influence in ways that produce favorable impacts on their stock prices?

Hypotheses

 Days with highly positive or negative social media sentiment will be associated with significant price changes (up or down) in the corresponding stock within a 1–3 day window.

Previous research

Over the past years, more research has looked at the intersection of behavioral finance and how it can affect the market. As traditional indicators are challenged with new and advance third party solutions, sentiment analysis has become a topic to research to possibly gain an advantage.

Social Media and Market Behavior

Bollen, Mao, and Zeng (2011) were among the first to show that Twitter (now X) feeds could predict changes in the Dow Jones Industrial Average with up to 86% accuracy. This study introduced social media as a factor in investor psychology.

Similarly, Sprenger et al. (2014) analyzed financial tweets and found that users on X who "providing above average investment advice", have more followers and engagement.

Reasoning to include previous research

For this paper, I found it important to include some previous studies. The reasoning behind it, is the strict rules and challenges with collecting and analyzing data.

Methodology

Data Collection

This study integrates three main types of data: social media posts, sentiment scores, and stock market performance. Social media posts and market performance was collected using open-source tools and public APIs to ensure reproducibility. Whereas sentiment scores were computed using VADER text analysis.

Social Media Data

Social media were gathered using Reddit, where posts were extracted using Praw, by creating a developer account and scraping subreddits for $TSLA. Posts were timestamped and indexed by ticker symbol, to make it easier for us to align it with data from the stock market.

Sentiment collection

Each post was processed through VADER sentiment analysis tool to assign a sentiment score. Sentiment scores were added to the data frame for analyzing, on a -1 (very negative) to +1 (very positive) scale.

Stock Price Data

Market performance was gathered using:

- Yahoo Finance API: Queried through the yfinance Python library for daily open, close, high, and low prices. (ended up not working, as the extraction rates were too low)
- Alpha Vantage: a third party service, that allowed for better extraction rates.

All data were merged together, to create a data frame that could be used to analyze the correlation between sentiment and market movements

Data Processing and Cleaning

To ensure the dataset gives the best results, n/a values were dropped, duplicates of posts were dropped to ensure no redundancy and rows that are missing key fields were dropped as well.

Visualization

To identify trends and patterns:

- Time series plots were generated showing sentiment scores against stock closing prices.

Analytical Approach

- Correlation between same day stock prices and sentiment
- Correlation between next day stock price and sentiment

Tools Used

- Python (Pandas, NumPy, NLTK, Matplotlib, Seaborn)
- Google Collab
- VADER and FINbert (for financial text understanding)
- Reddit Praw, Yahoo Finance, Alpha Vantage

Results & Visualizations

Sentiment vs. Stock Price Trends

To investigate the influence of social media sentiment on Tesla's stock performance, sentiment scores from Reddit posts mentioning "TESLA" were plotted against the company's daily closing stock price.

- Time Series Visualization

Tesla Price with High Sentiment Days
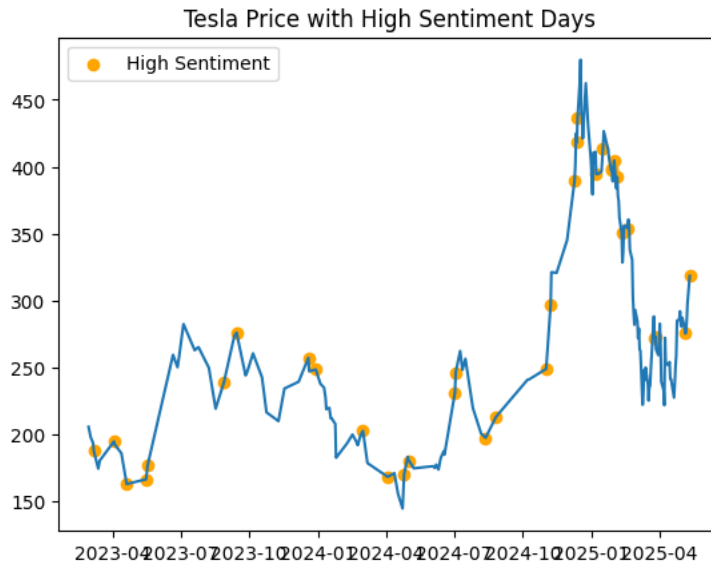
Days with high sentiment is marked as orange dots

Tesla Price with Low Sentiment Days
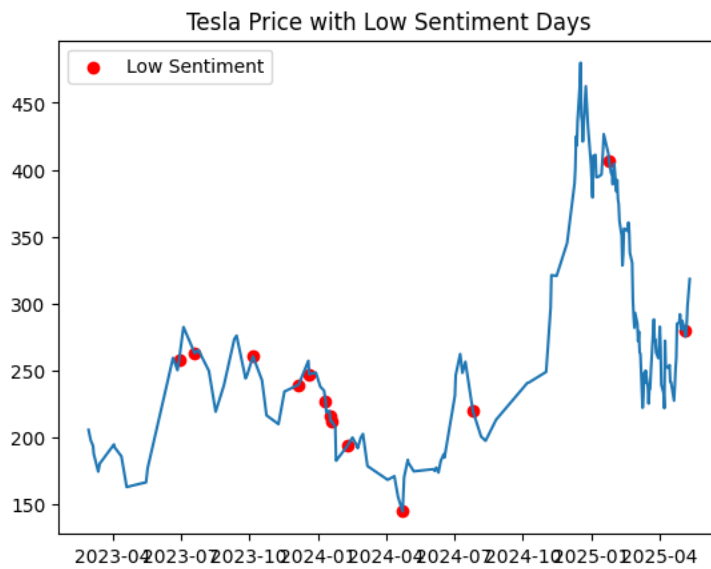
Days with low sentiment are marked as red dots

- Insights

  As we can see there is a lot of matching sentiment points to the way the stock price moves

Correlation Analysis (TSLA)

A Pearson correlation coefficient was calculated between the average daily sentiment score and the next day's price return.

- Result: The correlation was +0.12 on the same day and +0.15 on next days price, indicating a weak positive relationship.
- While statistically minor, this result suggests that high sentiment may be slightly associated with upward price movement, but the results are way to significant to be used as a reliable indicator.

Summary of Findings for Tesla

- Social media sentiment surrounding Tesla correlates weakly (+0.12 - +0.15) with short-term price returns.
- High or low sentiment days tend to be aligned with movements of the price, but the model was not able to predict which way.
- The sentiment-based model performs barely better than random, so this model would not be of good use.

Discussion

The analysis of Tesla's social media sentiment reveals a weak relationship between public sentiment and short-term market movements. With a Pearson correlation of +0.15 at best between sentiment scores and next-day returns, the findings suggest that while sentiment may provide some directional insight, it is far from reliable.

On the contrary, it may be included as a tool instead of price predictor. When sentiment hits extreme highs or lows, Tesla's stock is more likely to experience outsized movement the following day not necessarily up or down, but with larger-than-usual magnitude. This aligns with behavioral economics and finance, which shows that investors are action on emotion and letting

market reactions and social media activity alter their decisions, particularly for companies active in the media

From a strategic marketing perspective, these findings carry some reassurance. They may not be able to control social media, but at least regular social media posts are not pushing the price heavily. On the other hand, by strategically timing news and using Social media to spread them can be a different case.

However, in this study certain limitations were met. The computational power and limited access to API keys, made it difficult to generate the model needed for a task like this. There are a lot of good information, that can be taken from it, but overall the study shows that social media has very little impact on the movements of the stock market.

Future work should explore:

- Multi-ticker comparisons to test whether sentiment influence varies by company type.
- Integration of other features such as trading volume, earnings reports, and news headlines to enhance predictive modeling.

Despite these limitations, seeing that sentiment is high (either positive or negative) when the stock is moving. Could be a valuable lesson for a company to dig deeper into, especially if there are good news that needs to be spread quickly.

Conclusion

This project wanted to see whether social media sentiment can predict short-term stock movements and whether companies like Tesla can strategically leverage this in their marketing efforts. Through the integration of natural language processing and financial data analytics, the paper studied how posts from Reddit align with stock price changes for Tesla Inc. (TSLA).

The results show that there is a weak positive correlation between sentiment and next-day returns (+0.15), sentiment does not predict price direction. However, high sentiment were associated

with increased movements in the stock price. This could lead to opportunities for a marketing team to spread good news quick, or a warning to limit negative news to the best of their extent.

These findings shows that sentiment alone can not be traded for the old school financial analyses, but potentially could be a tool that is incorporated into a much bigger picture. Especially in marketing, investor relations, and risk management. Companies like Tesla, which has a lot of media exposure, monitoring sentiment trends and strategically timing announcing news could have an impact on the stock behavior.

While this study was limited by API access and computational power, it establishes a bridge from old school financial analyses into new and powerful tools that could make a difference. By expanding this idea across multiple stocks, applying more advanced NLP models and a better access to data, could offer better results. In the end the behavioral economics are always a factor in an investors mind, with everyone's opinions and financial "advice" one click away there will be some influence. In this case the hypothesis was not true, but with deeper research and better access there is a chance that the results could be more giving.

References

Ante, L. (2021). *How Elon Musk's Twitter activity moves cryptocurrency markets*. Finance Research Letters, 41, 101976.

https://www.sciencedirect.com/science/article/abs/pii/S187775031100007X?via%3Dihub

Bollen, J., Mao, H., & Zeng, X. (2011). *Twitter mood predicts the stock market*. Journal of Computational Science, 2(1), https://onlinelibrary.wiley.com/doi/10.1111/j.1468-036X.2013.12007.x