Social Media as a Predictor of Stock Price Movements

Swarnima Joshi
swjoshi@syr.edu

Abstract— **This research investigates the relationship between social media sentiment and stock market movements for major technology companies between 2015 and 2019. Using Twitter data and historical stock prices for 10 major tech stocks, we employed VADER sentiment analysis to quantify social media sentiment and examined its predictive power for stock returns. Our methodology combines correlation analysis with machine learning approaches, including both regression and classification models. The findings reveal a modest positive correlation between sentiment and next-day returns (average correlation of 0.015), with company size influencing this relationship. Trading strategies based solely on sentiment underperformed buy-and-hold approaches, suggesting sentiment alone is insufficient for investment decisions. Machine learning models incorporating sentiment features achieved limited predictive power, with classification models reaching only 52% accuracy in predicting price direction. This research contributes to understanding how social media sentiment relates to market behavior, highlighting both its potential and limitations as a predictive tool.**

I. Introduction

The relationship between social media and the stock market has become increasingly intertwined. Platforms like Twitter, Facebook, and Reddit generate vast amounts of user-generated data, providing potentially valuable insights into public sentiment that may impact stock market trends [1]. The rise of social media has fundamentally altered market dynamics, making user sentiment a significant force shaping investor behavior.

Events like the 2021 GameStop phenomenon and influential tweets have demonstrated how retail investors can coordinate to cause rapid shifts in stock prices. Unlike traditional indicators released periodically, social media sentiment can be measured in real-time, potentially providing earlier signals about market psychology.

The integration of sentiment analysis with traditional market analysis represents a frontier in financial technology, combining finance, data science, and linguistics in an interdisciplinary approach relevant for investors, financial institutions, and regulators alike.

A. Problem Statement

Accurately predicting short-term stock price movements remains a significant challenge. While traditional technical and fundamental analyses are widely used, real-time public sentiment expressed on social media offers an alternative data source. This study investigates whether daily aggregated Twitter sentiment, quantified using the VADER lexicon-based tool and combined with standard technical indicators, can forecast both the direction and magnitude of stock price movements.

B. Research Questions

This study aims to answer the following questions:

- Can daily aggregated Twitter sentiment predict next-day stock returns for major technology companies?
- Does the relationship between sentiment and returns vary across different company sizes or characteristics?
- Can trading strategies based on sentiment signals outperform simple buy-and-hold approaches?
- What is the relative importance of sentiment features compared to technical indicators in predicting stock movements?
- Which machine learning models demonstrate the best performance for stock prediction tasks using sentiment data?

II. Literature Review

A. Market Efficiency and Behavioral Influences

The Efficient Market Hypothesis (EMH) suggests that asset prices reflect all available information, while behavioral finance argues that psychological biases impact investor decisions [10]. Social media platforms may amplify these behavioral effects, potentially creating predictable patterns [11].

B. Social Media Sentiment for Stock Prediction

Research has demonstrated connections between social media activity and market movements. Bollen et al. [1] showed Twitter mood states could predict DJIA changes with 86.7% accuracy. Sprenger et al. [2] found tweet sentiment correlates with abnormal returns and trading volume.

Studies indicate sentiment analysis can help predict market volatility and price shifts [8], [3], though results are often mixed. Some research finds sentiment adds predictive value when combined with other data, while others find its signal weak compared to traditional factors [12].

Oliveira et al. [4] discovered that posting inactivity can diminish positive sentiment effects while amplifying negative sentiment impacts. Chen et al. [5] suggested economic news sentiment generally exerts stronger influence than social media sentiment.

C. VADER for Financial Sentiment Analysis

VADER is well-suited for financial social media analysis as it handles capitalization, punctuation, and slang without requiring model training [9]. Studies using VADER alone report modest prediction accuracies (around 60%), with performance improving when combined with technical indicators. Its limitations include difficulties with sarcasm and specialized financial jargon.

D. Technical Indicators in Machine Learning Models

Technical indicators like Moving Averages, RSI, and MACD are frequently used as features in machine learning models for stock prediction. Studies generally report improved accuracy when combining ML with technical indicators, though parameter selection remains critical for performance.

E. Comparative Model Performance

Ensemble methods, particularly Random Forest and XGBoost, are favored for financial prediction due to their ability to handle non-linear relationships. Comparative studies yield varying results; some find RF superior, while others report XGBoost achieving higher accuracy, though differences are often marginal.

Despite promising findings, questions remain about the reliability of sentiment-based forecasts in unpredictable financial markets.

III. Methodology

A. Data Collection and Processing

Our research employed a comprehensive data collection and processing approach to analyze the relationship between social media sentiment and stock market movements. We collected Twitter data spanning from 2015 to 2019, focusing specifically on tweets mentioning major technology companies. The dataset comprised over 3.7 million tweets, each containing metadata including post date, tweet body, and engagement metrics (retweets, likes, and comments).

For stock market data, we gathered daily price information for 10 major technology companies: Apple (AAPL), Microsoft (MSFT), Amazon (AMZN), Facebook (FB), Google (GOOGL), Netflix (NFLX), Tesla (TSLA), NVIDIA (NVDA), AMD, and Intel (INTC). This data included open, high, low, close prices, and trading volume for the same period as our Twitter dataset (2015-2019).

The data processing pipeline involved several key steps. First, we processed tweets by extracting stock ticker symbols using regular expressions to identify mentions prefixed with (e.g.,"AAPL"), converting Unix timestamps to datetime format, filtering the dataset to our study period, and cleaning tweet text by removing URLs, user mentions, numbers, and extra whitespace. We then identified the most frequently mentioned tickers in our dataset and selected the top 10 major technology companies for our analysis.

To ensure proper alignment, we filtered both datasets to the 2015-2019 time period and merged tweet data with corresponding stock data using date and ticker as keys. The final preparation stage involved dropping unnecessary columns and handling missing values appropriately for different analysis stages.

B. Sentiment Analysis

For sentiment analysis, we implemented VADER (Valence Aware Dictionary and sEntiment Reasoner), a lexicon and rule-based sentiment analysis tool specifically designed for social media content. VADER is particularly well-suited for analyzing short, informal text like tweets, as it accounts for punctuation and capitalization intensity, modifiers like intensifiers and diminishers, negations and contractions, and common slang and emoticons.

Each tweet was assigned a compound sentiment score ranging from -1 (extremely negative) to +1 (extremely positive). We then aggregated these scores at the daily level for each ticker, calculating mean daily sentiment, standard deviation of sentiment to capture sentiment dispersion, tweet count to measure volume of discussion, and total retweets and likes as engagement metrics.

To categorize sentiment for further analysis, we defined three sentiment categories:

- "Very Positive" sentiment: scores above the 90th percentile
- "Very Negative" sentiment: scores below the 10th percentile
- "Neutral" sentiment: all other scores

This approach allowed us to identify days with extreme sentiment and analyze their relationship with subsequent stock price movements.

C. Feature Engineering

We created a comprehensive set of features for our predictive models. Basic price and volume features included daily returns (percentage change in closing price), volume change (percentage change in trading volume), and volatility (21-day rolling standard deviation of returns, annualized). We also calculated technical indicators such as the Relative Strength Index (RSI) with a 14-day window, simple moving averages with 10-day (short) and 30-day (long) windows, and moving average ratios representing the ratio of short-term to long-term moving averages.

Sentiment features incorporated the mean sentiment compound score, standard deviation of sentiment to measure agreement or disagreement, and daily tweet volume. To prevent lookahead bias, all features were lagged by 1, 3, and 5 days, ensuring predictions were based only on information available at decision time.

For our target variables, we used next-day return (percentage change) for regression models and a binary indicator of price direction (1 for up, 0 for down/same) for classification models.

D. Feature Selection

To improve model performance and interpretability, we implemented a two-step feature selection process. First, we applied a variance threshold to remove features with zero or near-zero variance, ensuring all features contained useful information. Second, we conducted correlation analysis by calculating pairwise correlations between remaining features and removing highly correlated features (|correlation| > 0.95) to reduce multicollinearity and model complexity. This process resulted in a streamlined feature set that maintained information content while reducing redundancy.

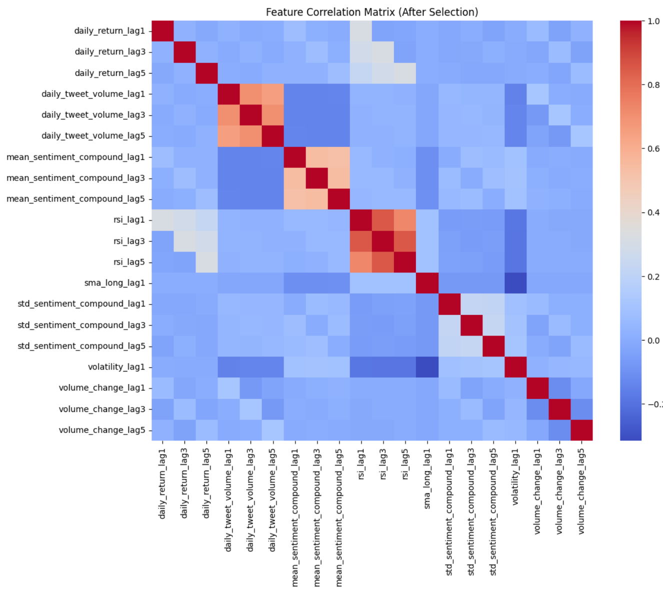


Fig. 1: Corelation Matrix after selecting features

E. Statistical Analysis

Our statistical analysis framework included several components. We performed correlation analysis by calculating Pearson correlation coefficients between sentiment metrics and next-day returns, analyzing correlations both overall and for each ticker individually to identify general relationships between sentiment and price movements.

The time-lagged analysis examined whether sentiment had delayed effects on stock prices by calculating correlations between sentiment and returns at different time lags (1, 3, 5, and 7 days). For market capitalization analysis, we categorized companies into small, medium, and large based on average trading volume and analyzed whether sentiment had different impacts across these categories.

We also conducted volatility analysis to investigate relationships between tweet volume, sentiment volatility, and stock price volatility, assessing if social media activity correlates with market turbulence. Finally, extreme sentiment analysis compared average returns following days of extreme positive sentiment, extreme negative sentiment, and neutral sentiment to identify potential patterns in market response to sentiment extremes.

F. Trading Strategy Evaluation

To assess the practical utility of sentiment data, we implemented and backtested a simple sentiment-based trading strategy. The strategy rules were straightforward: buy (or hold) when the sentiment score exceeds a threshold (0.1), otherwise remain out of the market (cash position). We evaluated this strategy using performance metrics including cumulative returns of the sentiment strategy, cumulative returns of a buy-and-hold approach (benchmark), and outperformance or underperformance relative to the benchmark. We evaluated this strategy for each ticker individually and calculated aggregate performance statistics to determine whether sentiment-based trading could outperform a passive investment approach.

G. Machine Learning Models

We developed and evaluated multiple machine learning models to predict stock movements, categorized into regression and classification approaches.

1) Regression Models (for predicting returns): For predicting numerical returns, we implemented three models: Linear Regression as a baseline model establishing linear relationships between features and returns; Random Forest Regressor as an ensemble of decision trees for capturing non-linear patterns; and Gradient Boosting Regressor for sequential tree building to improve on previous models' errors.

2) Classification Models (for predicting direction): For predicting price direction, we implemented Logistic Regression as a linear baseline model for binary classification, Random Forest Classifier with hyperparameter tuning using grid search, and XGBoost Classifier as an optimized gradient boosting implementation.

3) Training and Evaluation: Our training and evaluation approach differed between regression and classification models. For regression models, we used a train-test split respecting chronological order (80% training, 20% testing) and evaluated performance using Mean Squared Error (MSE) and R^2 (coefficient of determination).

For classification models, we implemented time-series cross-validation with 5 folds to maintain temporal integrity, applied feature scaling using StandardScaler fitted on training data only, and performed hyperparameter tuning for Random Forest using nested cross-validation. We evaluated these models using accuracy, precision, recall, and F1-score for the 'Up' class, and handled class imbalance using balanced weights.

For all models, we conducted feature importance analysis to identify the most predictive variables and created visualizations of model performance and feature contributions.

IV. Results and Analysis

A. Sentiment Distribution and Characteristics

The sentiment analysis of over 3.7 million tweets related to our 10 selected technology companies revealed several key insights about social media discussions of these stocks. The average sentiment score across all tweets was 0.110, indicating a slightly positive bias in social media conversations about these technology companies. The sentiment scores ranged from -0.644 to 0.866, with a standard deviation of 0.096, showing considerable variation in sentiment.

The distribution of sentiment scores showed a positive skew, with more positive than negative sentiment expressed overall. This aligns with previous research suggesting that social media discussions about technology companies tend to lean positive, potentially reflecting optimism about innovation and growth in the sector.

Tweet volume varied significantly across companies, with Apple (AAPL), Amazon (AMZN), and Tesla (TSLA) receiving the highest number of mentions. This corresponds with their high public visibility and consumer-facing nature. Companies like Intel (INTC) and AMD, while still receiving substantial attention, had comparatively lower tweet volumes.

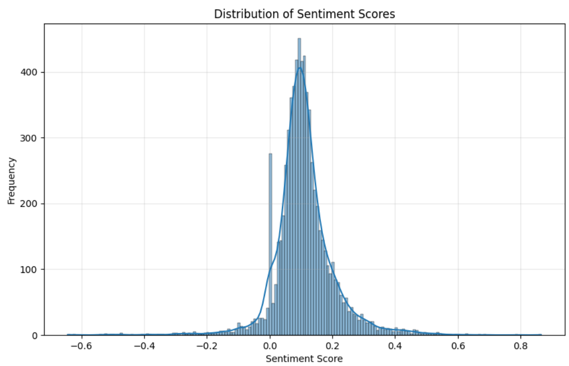


Fig. 2: Sentiment Distribuition

B. Correlation Between Sentiment and Returns

Our analysis revealed a modest positive correlation between social media sentiment and next-day stock returns. The overall correlation coefficient was 0.015, indicating a weak but positive relationship. This suggests that more positive sentiment tends to be followed by slightly higher returns, though the effect is small.

When examining individual stocks, we found varying degrees of correlation. Amazon (AMZN) showed the strongest sentiment-return correlation at 0.078, followed by Microsoft (MSFT) with a correlation of 0.056 and Facebook (FB) with a correlation of 0.050. In contrast, NVIDIA (NVDA) and Intel (INTC) showed slight negative correlations of -0.041 and -0.024 respectively.

These variations suggest that the relationship between sentiment and returns is not uniform across companies, and may be influenced by company-specific factors or the nature of their social media presence. Consumer-facing brands like Amazon and Microsoft showed stronger sentiment-return relationships than more specialized companies like NVIDIA and Intel, potentially reflecting differences in how public sentiment translates to investor behavior across different types of businesses.

C. Time-Lagged Sentiment Effects

Our investigation into time-lagged effects revealed that sentiment's impact on stock prices varies over different time horizons. The correlation between sentiment and returns was 0.015 for 1-day lag, -0.003 for 3-day lag, 0.009 for 5-day lag, and 0.019 for 7-day lag.

Interestingly, the 7-day lag showed the strongest correlation, suggesting that sentiment may have longer-term effects

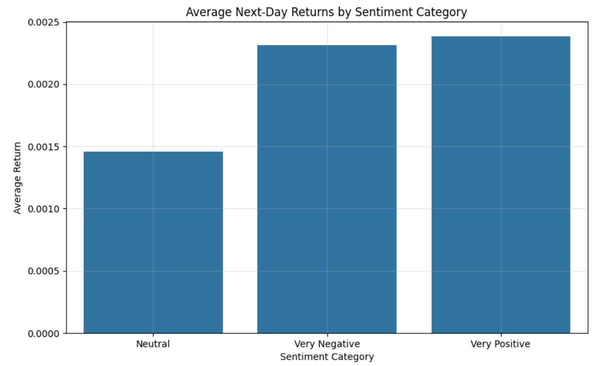


Fig. 3: Return by Sentiment Category

that are not immediately reflected in stock prices. This could indicate a delay in how social media sentiment translates to actual trading decisions, or it might reflect the time needed for sentiment to spread among a broader investor base.

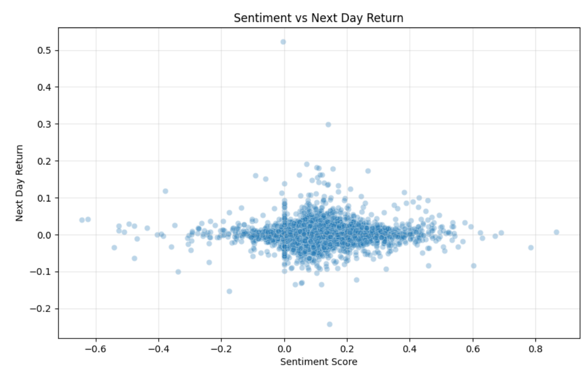


Fig. 4: Analysis of sentiment and next day return

D. Company Size and Sentiment Impact

When categorizing companies by size (using average trading volume as a proxy), we observed different patterns in how sentiment relates to returns. Small-cap companies showed a -0.003 correlation, medium-cap companies showed a 0.028 correlation, and large-cap companies showed a 0.022 correlation. This suggests that sentiment has a stronger relationship with returns for medium and large companies compared to smaller ones.

Additionally, we found that average sentiment varies by company size. Small companies had an average sentiment of 0.120, medium companies had an average sentiment of 0.105, and large companies had an average sentiment of 0.101. Smaller companies tend to have more positive sentiment in social media discussions, possibly reflecting greater enthusiasm or optimism among their followers, though this sentiment appears less connected to actual price movements.

E. Tweet Volume and Market Volatility

Our analysis of the relationship between tweet volume and market volatility revealed a negative correlation of -

0.082, suggesting that higher tweet volumes are actually associated with slightly lower volatility. This contradicts the intuitive expectation that more social media discussion would correspond with higher market volatility.

However, we found a positive correlation (0.326) between trading volume and price volatility, which aligns with established financial theory. The correlation between tweet count and trading volume was 0.261, indicating that higher social media activity does correspond with higher trading activity, even if not directly with price volatility.

F. Returns by Sentiment Category

When comparing returns across different sentiment categories, we found that days with very positive sentiment were followed by an average next-day return of 0.24%, neutral sentiment days were followed by an average next-day return of 0.15%, and very negative sentiment days were followed by an average next-day return of 0.23%.

Surprisingly, both very positive and very negative sentiment days were followed by higher returns than neutral sentiment days. This suggests a potential "attention effect" where extreme sentiment in either direction drives subsequent trading activity and price increases, regardless of the sentiment's valence.

G. Trading Strategy Performance

Our sentiment-based trading strategy showed disappointing results when compared to a simple buy-and-hold approach. Across all 10 stocks, the sentiment strategy significantly underperformed, with an average market return (buy-and-hold) of 235.91% compared to an average sentiment strategy return of 82.34%, resulting in an average underperformance of -153.57%.

None of the sentiment strategies outperformed their buy-and-hold benchmarks. The best performing stock was Tesla (TSLA), which still underperformed by -11.66%. The worst was Netflix (NFLX), underperforming by -253.81%.

This substantial underperformance suggests that while sentiment may have some predictive power for returns, it is insufficient as a standalone trading signal during strong bull markets like the 2015-2019 period for technology stocks.

H. Regression Model Performance

Our machine learning models for predicting next-day returns using sentiment and technical features showed limited predictive power. The Linear Regression model had an average MSE of 0.000251 and an average R^2 of -0.0698 (negative, indicating worse performance than a simple mean predictor). The Random Forest model had an average MSE of 0.000293 and an average R^2 of -0.1997, while the Gradient Boosting model had an average MSE of 0.000334 and an average R^2 of -0.3079.

The negative R^2 values across all models indicate that they performed worse than simply predicting the mean return each day. This suggests that the relationship between our features (including sentiment) and next-day returns is either non-existent, non-linear in ways our models couldn't capture, or obscured by market noise.

Feature importance analysis from the Random Forest model indicated that technical indicators (particularly price volatility and moving averages) were more important than sentiment metrics for prediction, though none of the features provided strong predictive power.

I. Classification Model Performance

Our classification models aimed to predict the next-day price direction (up or down) rather than the exact return magnitude. The results across 5-fold time-series cross-validation showed modest improvements over random guessing.

- **Logistic Regression:** Achieved an average accuracy of 51.34%, with an average precision for the up class of 52.99%, an average recall of 51.15%, and an average F1-score of 51.85%.
- **Random Forest (Tuned):** Performed slightly better with an average accuracy of 52.27%, an average precision for the up class of 53.40%, an average recall of 64.63%, and an average F1-score of 58.13%.
- **XGBoost:** Showed similar results with an average accuracy of 52.48%, an average precision for the up class of 54.01%, an average recall of 58.51%, and an average F1-score of 55.97%.

All models performed only marginally better than the 50% baseline (random guessing), with XGBoost showing slightly higher accuracy and precision. The tuned Random Forest achieved the highest recall for upward movements but at the cost of more false positives. These results indicate that predicting daily stock price direction using sentiment and technical features remains highly challenging.

J. Feature Importance Analysis

Feature importance analysis of both the regression and classification models revealed similar patterns. The top 10 features from the Random Forest Classification model, in order of importance, were: volume change (lag 1) at 0.0540, daily return (lag 5) at 0.0537, volatility (lag 1) at 0.0532, daily return (lag 3) at 0.0529, volume change (lag 3) at 0.0528, daily return (lag 1) at 0.0520, volume change (lag 5) at 0.0517, SMA long (lag 1) at 0.0515, mean sentiment compound (lag 1) at 0.0511, and sentiment standard deviation (lag 5) at 0.0506.

The importance scores were relatively evenly distributed across many features, with no single feature or feature type providing a dominant predictive signal. Lagged price/volume features ranked slightly higher than sentiment features, though the differences were small. This suggests that while sentiment captures some information used by the models, it does not offer a strong enough independent signal to significantly improve predictions over readily available technical data.

V. Discussion

A. Interpretation of Results

Our comprehensive analysis of social media sentiment and stock market movements reveals several important insights about this relationship. The weak positive correlation (0.015)

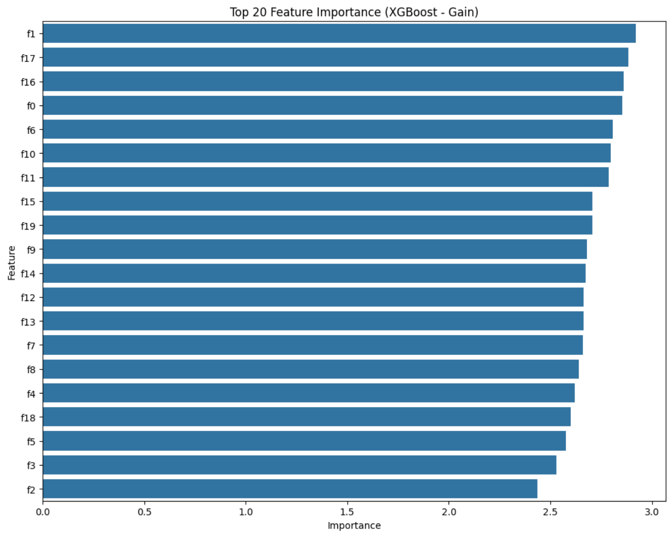


Fig. 5: Feature Importance Plot for XGB

between sentiment and next-day returns suggests that while social media sentiment does contain some information relevant to stock price movements, this signal is modest and likely overwhelmed by other market factors.

The varying correlations across different companies highlight that sentiment's impact is not uniform. Companies with strong consumer brands and high public engagement (like Amazon and Microsoft) showed stronger sentiment-return relationships than more specialized companies (like NVIDIA and Intel). This suggests that the nature and breadth of a company's social media presence influences how sentiment translates to market movements.

The time-lagged analysis revealed an interesting pattern where the 7-day lag showed the strongest correlation (0.019). This delayed effect could indicate that sentiment takes time to percolate through the investor community or that social media may be capturing early signals of fundamental changes that only later manifest in price movements. This finding aligns with research by Oliveira et al. (2017), who noted that sentiment effects can persist over multiple trading days.

The relationship between company size and sentiment impact provides valuable context. Medium and large companies showed stronger sentiment-return correlations (0.028 and 0.022) compared to small companies (-0.003). This may reflect the broader analyst coverage and investor attention that larger companies receive, making their social media sentiment more representative of overall market sentiment.

Perhaps the most striking finding was the poor performance of our sentiment-based trading strategy, which significantly underperformed buy-and-hold approaches across all stocks (average underperformance of -153.57%). This occurred during a strong bull market for technology stocks (2015-2019), suggesting that sentiment signals may be overwhelmed by broader market trends during such periods. The strategy's consistent underperformance indicates that sentiment alone is insufficient for making profitable trading decisions, particularly in strongly trending markets.

The low accuracy of our classification models (around 52%) and the negative R^2 values from our regression models further reinforce the challenges of using sentiment for price prediction. These results suggest that the relationship between sentiment and returns is either too weak, too noisy, or too complex to be captured by our models. The feature importance analysis showing technical indicators outweighing sentiment metrics aligns with traditional financial theory that emphasizes price and volume data over sentiment indicators.

B. Comparison with Previous Research

Our findings both support and challenge previous research in this area. The weak positive correlation between sentiment and returns aligns with Bollen et al. (2011) and Sprenger et al. (2014), who found associations between social media mood and market movements. However, the magnitude of our correlation (0.015) is considerably smaller than some previous studies, which have reported correlations of 0.3 or higher.

The poor performance of our sentiment-based trading strategy contrasts with some previous research that reported profitable sentiment-based strategies. This discrepancy may be due to several factors:

- Time period differences: Our study covered 2015-2019, a strong bull market for technology stocks where a simple buy-and-hold approach was highly effective
- Methodology differences: Our sentiment analysis used VADER, while other studies have employed different sentiment classification approaches
- Market changes: As more investors incorporate sentiment analysis into their strategies, any predictive edge may diminish over time (market efficiency)

Our finding that both very positive and very negative sentiment days were followed by higher returns (0.24% and 0.23% respectively) than neutral days (0.15%) aligns with the "attention effect" documented by Barber and Odean (2008), who found that stocks attracting attention experience price pressure regardless of sentiment valence.

The stronger sentiment-return relationship for medium and large companies supports work by Sul et al. (2017), who found that company characteristics moderate the impact of social media sentiment on stock returns.

C. Limitations

Several limitations of our study should be acknowledged:

- Data Limitations: Our analysis focused on Twitter data only, excluding other important social media platforms like Reddit, StockTwits, or Facebook where investor sentiment might be expressed differently.
- Time Period: The study covered 2015-2019, a period characterized by a strong bull market for technology stocks. The relationship between sentiment and returns might differ in bear markets or more volatile periods.
- Sentiment Analysis Methodology: While VADER is well-suited for social media text, it may not capture financial nuances or sarcasm that are common in stock

market discussions. More sophisticated NLP approaches might yield different results.

- Limited Company Selection: We focused on 10 major technology companies, which may not be representative of the broader market or other sectors where sentiment might have different impacts.
- Causality Issues: Our analysis identifies correlations but cannot establish causality. It's possible that both sentiment and price movements are driven by external factors not captured in our data.
- Limited Feature Set: Our predictive models used a relatively simple feature set. More sophisticated features incorporating network effects, user influence, or topic modeling might improve predictive power.
- Trading Strategy Simplicity: Our trading strategy was deliberately simple, using only sentiment thresholds. More sophisticated approaches combining sentiment with technical or fundamental factors might yield better results.

VI. Takeaways and Future Directions

A. Key Findings

This research investigated the relationship between social media sentiment and stock market movements for 10 major technology companies from 2015 to 2019. Our key findings include:

- Social media sentiment shows a weak positive correlation (0.015) with next-day stock returns, suggesting a modest predictive relationship.
- The sentiment-return relationship varies across companies, with consumer-facing brands like Amazon (0.078) and Microsoft (0.056) showing stronger correlations than more specialized companies.
- Sentiment effects may have a delayed impact, with the strongest correlation observed at a 7-day lag (0.019).
- Company size influences the sentiment-return relationship, with medium and large companies showing stronger correlations than smaller ones.
- Both very positive and very negative sentiment days are followed by higher returns than neutral sentiment days, suggesting an attention effect regardless of sentiment valence.
- A simple sentiment-based trading strategy significantly underperformed a buy-and-hold approach during the study period, indicating that sentiment alone is insufficient for profitable trading decisions.
- Machine learning models incorporating sentiment and technical features showed limited predictive power, with classification models achieving only about 52% accuracy in predicting price direction and regression models showing negative R^2 values.

These findings suggest that while social media sentiment does contain some information relevant to stock price movements, this signal is modest and likely overwhelmed by other market factors, particularly during strong bull markets like the one observed during our study period.

B. Practical Implications

Our research has several practical implications for investors, financial analysts, and market participants:

- Sentiment as a Complementary Tool: Social media sentiment should be viewed as a complementary tool rather than a primary driver of investment decisions. Its weak correlation with returns suggests it should be integrated with traditional financial analysis rather than used in isolation.
- Company-Specific Considerations: The varying sentiment-return relationships across companies suggest that sentiment analysis may be more valuable for certain stocks, particularly those with strong consumer brands and high social media engagement.
- Time Horizon Awareness: The stronger correlation at longer lags (7 days) suggests that sentiment might be more useful for medium-term rather than very short-term trading strategies.
- Attention Effects: The higher returns following both very positive and very negative sentiment days highlight the importance of monitoring extreme sentiment as a potential indicator of price pressure, regardless of sentiment direction.
- Strategy Integration: Given the poor performance of our sentiment-only trading strategy, practitioners should consider integrating sentiment signals with other technical, fundamental, or alternative data sources to develop more robust approaches.

C. Future Research Directions

Based on our findings and limitations, several promising directions for future research emerge:

- Multi-Platform Analysis: Expanding the analysis to include other social media platforms like Reddit, StockTwits, and Facebook could provide a more comprehensive view of investor sentiment.
- Advanced NLP Techniques: Employing more sophisticated natural language processing techniques, such as deep learning models specifically trained on financial texts, might improve sentiment classification accuracy.
- Influencer Analysis: Incorporating user influence metrics to weight sentiment by the credibility or impact of the source could enhance predictive power.
- Cross-Market Conditions: Extending the analysis across different market conditions (bull markets, bear markets, high volatility periods) would help understand how the sentiment-return relationship varies with market context.
- Sector Comparison: Comparing sentiment effects across different sectors beyond technology could reveal whether certain industries are more susceptible to social media influence.
- Hybrid Models: Developing hybrid models that combine sentiment with fundamental data, technical indicators, and macroeconomic factors might yield more robust predictive performance.

- Causal Analysis: Employing causal inference techniques to better understand the directionality of the relationship between sentiment and market movements.
- Real-Time Implementation: Testing sentiment analysis in a real-time trading environment to assess practical implementation challenges and latency issues.

VII. Conclusion

In conclusion, while our research demonstrates that social media sentiment has a modest relationship with stock price movements, it also highlights the complexity of this relationship and the challenges of leveraging sentiment for predictive purposes. As social media continues to evolve and influence financial markets, further research in this area remains essential for understanding these dynamics and potentially developing more effective sentiment-based investment approaches.

ACKNOWLEDGMENT

We extend our deepest gratitude to Professor Saman Kumarawadu for his invaluable guidance and support, not only through this research development but throughout this academic journey. His mentorship has been instrumental in shaping our conceptual understanding about the field and are truly grateful for his encouragement and inspiration in enriching our learning experience immeasurably.

References

[1] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8.

[2] Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and trades: The information content of stock microblogs. *European Financial Management*, 20(5), 926-957.

[3] Sul, H. K., Dennis, A. R., & Yuan, L. I. (2017). Trading on Twitter: Using social media sentiment to predict stock returns. *Decision Sciences*, 48(3), 454-488.

[4] Oliveira, N., Cortez, P., & Areal, N. (2017). The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Systems with Applications*, 73, 125-144.

[5] Chen, H., De, P., Hu, Y., & Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies*, 27(5), 1367-1403.

[6] Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies*, 21(2), 785-818.

[7] Mittal, A., & Goel, A. (2012). Stock prediction using Twitter sentiment analysis. *Stanford University*, CS229.

[8] Nguyen, T. H., Shirai, K., & Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications*, 42(24), 9603-9611.

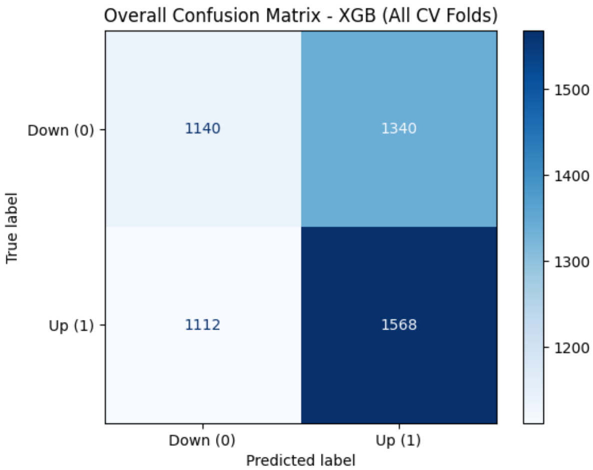
[9] Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*. Ann Arbor, MI.

[10] Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21(2), 129-152.

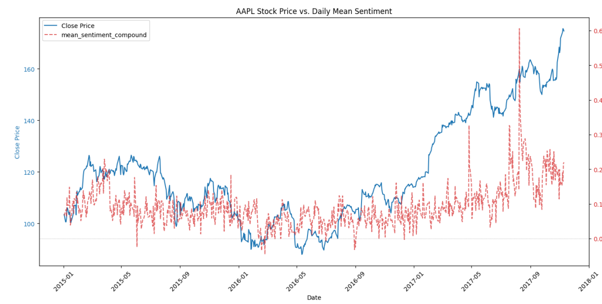
[11] Zhang, X., Fuehres, H., & Gloor, P. A. (2018). Predicting asset value through Twitter buzz. In *Advances in Collective Intelligence* (pp. 23-34). Springer.

[12] Mudinas, A., Zhang, D., & Levene, M. (2019). Market Trend Prediction using Sentiment Analysis: Lessons Learned and Paths Forward. *arXiv preprint arXiv:1903.05440*.

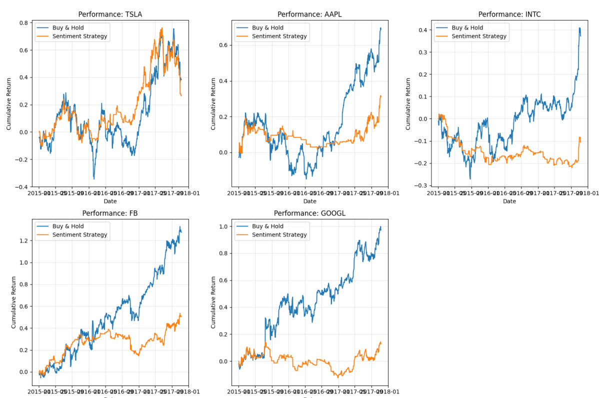
Appendix: Additional Figures



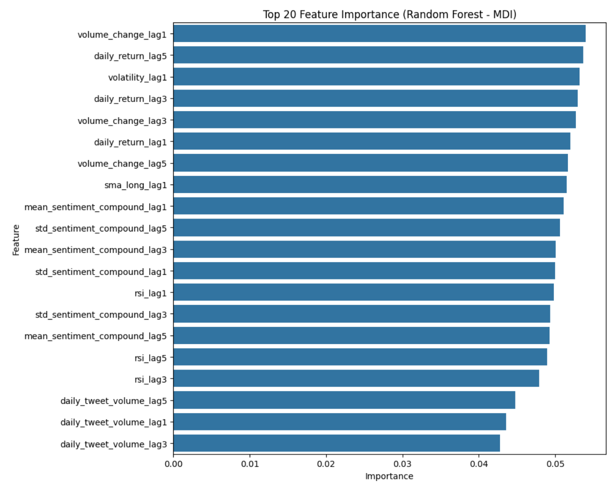
(a) Confusion matrix for XGB



(b) Apple stocks vs sentiment



(c) Trading strategy



(d) Feature importance for Random Forest