Can public sentiment predict Stock returns?

Final Project: Data Science for Finance

The Wharton School

Professor Roberts

Ali Khambati, Jackson Joffe, Namedia Ozingbe, and Donn Boddie

Background & Data Collection

Background

- The explosion of retail traders and the rise of social media has created a paradigm shift in stock movement
- Public sentiment influences retail traders who consist of a majority of daily volume
- We attempt to quantify the exact effect of public sentiment on stock returns

Data Sources

- Twitter; python package TweetCollector used to collect daily tweets mentioning Apple, Microsoft, Tesla, Google, and Amazon
- WRDS; collected stock price, volume and overall market data

Selected Initial Visualizations

- Sentiment is normally distributed
- Outliers correspond to sporadic events like earnings news or leadership transitions

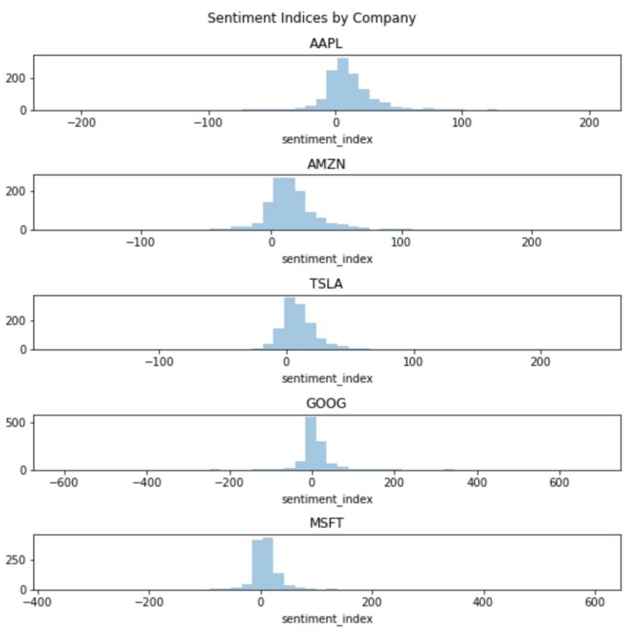

Figure 2: Distributions of Sentiment Scores

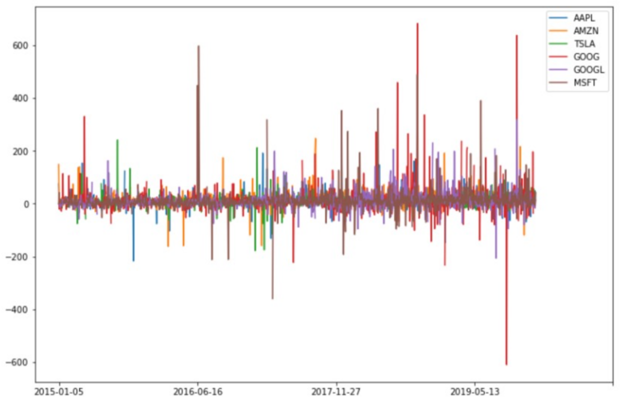

Figure 1: Series of Sentiment Scores from 2015-2020

Random Forest Regression

- R^2: 88% over training data
    - High evidence of overfitting – poor cross validation and test data
- Evidence of weak correlation with public sentiment; indicates that social media is a proxy for market information but does not influence ultimate capital deployment decisions
- Asset pricing is more complex

Conclusion

Initial hypothesis that sentiment is highly correlated with daily returns is incorrect, but evidence suggests a weak positive correlation

Limited support for herd mentality in the context of trading volume and daily returns

Thin volume stocks are most susceptible to sentiment-based volatility, while institutional interest in blue chip companies shield, to an extent, retail volatility

References

- Aboody, D., Even-Tov, O., Lehavy, R., Trueman, B. (2018). Overnight returns and firm-specific investor sentiment. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2554010
- Bollen, J., Mao, H. and Zeng, X., 2011. Twitter mood predicts the stock market. Journal of Computational Science, [online] 2(1), pp.1-8. Available at: <https://arxiv.org/abs/1010.3003> [Accessed 2 December 2021].
- Chen, Y., Zhao, H., Li, Z., Lu, J. (2020). A dynamic analysis of the relationship between investor sentiment and stock market realized volatility: Evidence from China. PLOS ONE, 15(12). https://doi.org/10.1371/journal.pone.0243080
- C.J. Hutto, vaderSentiment, (2021), GitHub repository, https://github.com/cjhutto/vaderSentiment
- Dobbs, R., Huyett, B. and Koller, T., 2021. The CEO’s guide to corporate finance. [online] mckinsey.com. Available at: <https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/the-ceos-guide-to-corporate-finance> [Accessed 2 December 2021].
- Fama, E., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, [online] 25(2). Available at: <https://www.jstor.org/stable/2325486> [Accessed 2 December 2021].
- Leask, H., 2021. Hedge funds’ use of alternative data tipped to surge, new industry study finds. [online] Hedgeweek. Available at: <https://www.hedgeweek.com/2020/05/04/285283/hedge-funds-use-alternative-data-tipped-surge-new-industry-study-finds> [Accessed 2 December 2021].
- Moldovan, E., 2021. Investors Psychology And The Herd Effect On The Financial Markets. [online] Ideas.repec.org. Available at: <https://ideas.repec.org/a/aio/rteyej/v1y2010i15sp21-26.html> [Accessed 2 December 2021].
- Omer Metin, TweetCollector, (2020), GitHub repository, https://github.com/omer-metin/TweetCollector
- Perrin, A. and Atske, S., 2021. About three-in-ten U.S. adults say they are ‘almost constantly’ online. [online] Pew Research Center. Available at: <https://www.pewresearch.org/fact-tank/2021/03/26/about-three-in-ten-u-s-adults-say-they-are-almost-constantly-online/> [Accessed 2 December 2021].
- Statista. 2021. U.S. Twitter reach by age group 2021 | Statista. [online] Available at: <https://www.statista.com/statistics/265647/share-of-us-internet-users-who-use-twitter-by-age-group/> [Accessed 2 December 2021].
- Wang, G., Yu, G., Shen, X. (2020). The effect of online investor sentiment on Stock Movements: An LSTM approach. Complexity, 2020, 1–11. https://doi.org/10.1155/2020/4754025


Thank You
Wharton
UNIVERSITY of PENNSYLVANIA