Earnings Report Project: Written Report

Allyssa Utecht, Jimmy Horvath, and Andrew Plate

For the Earnings Report Project, our group began by brainstorming approaches to find indicators in market reactions based on highlights from the MD&A. After reviewing the content we reviewed for our class and using our best judgment, our group decided that the sentiment analysis of sentences would be our leading indicator. Within the MD&A, we wanted to analyze the first three sentences and the last sentence. Our research found that “it was not the sentiment of the annual report itself, but the change in sentiment that was most important”. [[1]](#footnote-1) Our group felt that the introduction and conclusion to the document would hold more weight than the middle, and could pick up on any change in sentiment from beginning to end.

Furthermore, we wanted to implement another aspect of sentiment analysis with a lexicon of single words. Our group believed we could find a set of words, both positive and negative, that could be picked out of the MD&A and correlate with earnings persistence, as “the transparency of a lexicon-based approach represents a significant advantage in fields such as economics and finance, where stakeholders favor the clear interpretability of the methods and outcomes.”[[2]](#footnote-2) Additionally, if implemented correctly, a single-word lexicon can remove human interpretation to give a score solely based on the words alone. Next, our group decided the fog index would be an essential indicator to include within our model through conducting research and reviewing the second homework for class. MD&As that are "foggier" or harder to read are less persistent with their earnings, whether the author is doing this on purpose.[[3]](#footnote-3) The latter aspect is crucial, as it allows our methodology to work without human bias. Whether the author is bogging down the document on purpose or solely attempting to explain poor performance better, the model would read it with the same negativity or positivity. Our group also wanted to include the firm and market performance five days prior to the release of the MD&A. This would give us the best real-time indicator of our desired firm's performance and future outlook.

After brainstorming various indicators and building our primary model, with the help of Professor Davis’ model, our group realized adjustments needed to be made with the sentiment analysis of the MD&A. With a suggestion from Professor Davis, we realized that not all MD&A’s had a third sentence to analyze, returning an error when the code was run. We fixed this by adding a “try” and “pass” function to negate any MD&A that does not have a third sentence, allowing the code to keep running.

Next, our group realized the need for bigrams within our lexicon-based sentiment analysis. After digging through various sample MD&As, we found that most contain the word “COVID-19”, providing a strong indicator of a negative year. If firms are doing poorly in a global downturn, our group considers them more likely to use the event as an excuse for financial failures. On the contrary, firms performing financially well in these times are more likely to focus on proudly and perhaps overconfidently providing evidence for firm success despite global challenges, in opposition to mentioning how “COVID-19” impacted business. The phrases “strategic advancement” and “increased revenue” were also added, because while specific, provide strong indicators of positive earnings. We also tweaked the weights of the sentiment analysis and fog index of our code to better match their indicators.

Our final model began by setting the score to a baseline of zero. The subsequent functions either added to or subtracted from that score, based on a variety of predictors. First, we thought the first three sentences and the last sentence of the MD&A carried much weight towards an indicator of earnings persistence. The tokenization of the first sentence is multiplied by three, because it most likely offers a summary of the MD&A. Most standard writing structures include topic sentences to introduce what is to follow, carrying more weight than the following sentences. The second sentence is multiplied by two for the same reason, but holds less weight than the first sentence. The scores of the third and last sentences are not multiplied, as they offer insight into the MD&A but do not hold as much weight as the first couple sentences. Because of this, a fog index would be able to analyze the first sentence and its clarity. If the first sentence is “foggy” and indicates the rest of the MD&A to be negative, then the function will correlate it as such through the increased weight in the function.

The next section of code adjusts the score depending on the firm and market performance of the previous five days. These lines are run under ‘if’ statements to accept the possibility of market performance and firm performance being above or below zero. These scores are multiplied by one-hundred because the performance is a very small number. Removing some of the decimals was necessary to emphasize the impact of prior performance on future market reaction.

Finally, the last section of text analysis is the lexicon tokenizer (analysis). In the positive list, there are twenty-nine words, while in the negative there are twenty-seven words. Additionally, two positive phrases and one negative phrase are added to the lexicon. These positive words and phrases are multiplied by two, while their negative counterparts are multiplied by three. We chose these factors due to negativity bias, where humans commonly place more significance on negative events rather than on positive ones[[4]](#footnote-4).

Looking forward, when our model is tested on MD&As that we have not seen before, we expect our model to perform relatively well. It was imperfect in predicting market reaction on the models we had already seen. Although, it was overall more correct than not, so we expect it to perform similarly on new text. Perhaps our model would struggle identifying companies with vague topic sentences or introductions having nothing to do with firm performance. Thus, it truly depends on which MD&As the model analyzes and how they are formatted.

Bibliography

Consoli, S., Barbaglia, L., & Manzan, S. (2022). Fine-grained, aspect-based sentiment analysis on economic and financial lexicon. *Knowledge-Based Systems*, 108781. <https://doi.org/10.1016/j.knosys.2022.108781>

de Souza, J. A. S., Rissatti, J. C., Rover, S., & Borba, J. A. (2019). The linguistic complexities of narrative accounting disclosure on financial statements: An analysis based on readability characteristics. *Research in International Business and Finance*, *48*, 59–74. <https://doi.org/10.1016/j.ribaf.2018.12.008>

Hajek, P., Olej, V., & Myskova, R. (2013, October). *Forecasting Stock Prices using Sentiment Information in Annual Reports – A Neural Network and Support Vector Regression Approach* [Review of *Forecasting Stock Prices using Sentiment Information in Annual Reports – A Neural Network and Support Vector Regression Approach*]. WSEAS TRANSACTIONS on BUSINESS and ECONOMICS. [https://www.wseas.org/multimedia/journals/economics/2013/235702-202.pd](https://www.wseas.org/multimedia/journals/economics/2013/235702-202.pdf)f

Henry, E. (2006). Market Reaction to Verbal Components of Earnings Press Releases: Event Study Using a Predictive Algorithm. *Journal of Emerging Technologies in Accounting*, *3*(1), 1–19. <https://doi.org/10.2308/jeta.2006.3.1.1>

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, *45*(2-3), 221–247. <https://doi.org/10.1016/j.jacceco.2008.02.003>

OpenAI. (2023). *ChatGPT* (November 10 version) [Large language model].

Yekini, L. S., Wisniewski, T. P., & Millo, Y. (2016). Market reaction to the positiveness of annual report narratives. *The British Accounting Review*, *48*(4), 415–430. <https://doi.org/10.1016/j.bar.2015.12.001>

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. <https://doi.org/10.1037/0033-2909.134.3.383>

Zhong, N., & Ren, J. (2022). Using sentiment analysis to study the relationship between subjective expression in financial reports and company performance. *Frontiers in Psychology*, *13*. <https://doi.org/10.3389/fpsyg.2022.949881>

1. https://www.wseas.com/journals/bae/2013/235702-202.pdf [↑](#footnote-ref-1)
2. https://www.sciencedirect.com/science/article/pii/S0950705122003677 [↑](#footnote-ref-2)
3. https://www.sciencedirect.com/science/article/pii/S0165410108000141 [↑](#footnote-ref-3)
4. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3652533/ [↑](#footnote-ref-4)