

Problem Statement:

By analyzing the earning call transcripts, would the classification model provide value for investors to trade next day?

Data Source:

1. Text sources: Earning call transcripts in the major automobile maker such as Ford, General Motors, Tesla, Fiat Chrysler are retrieved from seekingalpha.com with essential subscription.
2. Numerical sources: Stock price and volumes are retrieved from Yahoo Finance

Modeling Objective:

1. While the call transcripts are released, this model will predict either the stock price will be up or down next day.
2. Since most Earning per share and quarter conference calls will be held after the market closes or occasionally before the market. Investors have a short period of time to analyze how the fluctuation of stock price will be before trade. The value of adapting this model will provide buy or sell signal to the speculative investors, confirming with their intuition on the trade.

EDA and Modeling Process

1. After merging the text and numerical dataset, transcripts will be splitted into Management Discussion (DA) and Analyst Question and Answer (QA). Features engineering will be derived from these two sessions. New features such as syntax sentiment, polarity, subjectivity, number of difficult words, readability indices, etc. will be added. After all, the numerical columns will be part of the data analysis
2. Second, text alone will be vectorized with n-grams for a new modeling process. Features of importance could provide investors what phrase or words matters
3. When the generalized model is completed, confusion matrix will be adapted to examine the performance of the classification. Additionally, a specialized model that only included dataset for one company will be modeled, followed by above engineering process. This aims to audit whether the classification model may provide a better prediction in one company rather than the entire industry.
4. Lastly, new stop words will be collected and introduced to the model for another round of evaluation.
5. The final adaption with XGBoost will provide good execution spend and model performance. While reviewing the bigrams for feature of importance, surprisingly, entity, location, and name are excluded. Generalized model will be defeated by noisy since each company has unique aspects and business operation, even though normalize the scale of the data set. However, the use of unsupervised learning such as KNN may be not applicable in this case.

Recommendation

1. The implication of using this model could help investors to understand of tone and language complexity of the management team and therefore suggest the market is either go up or down next day.
2. Investors could receive up or down signal from the classification model. They could compare whether these signals will be helpful to validate their intuition on next day trading when the earning call is released.
3. In general, the EDA process and model result suggest that firms with a high readability index tend to miss the target, suggesting a possibility that management team tries to increase their language complexity to avoid taking some issues. Therefore, investors could refer to the change of the readability index during the conference calls. While confirming with the numerical model, readability and negative sentiment by Analyst's QA are weighted heavily on coefficients.
4. Automobile industry focuses on the expansion of North America and China, Inventory, margin profit, and future plan are the importance features. These topics are the important features suggested by the XG Boost model.

Limitations

1. More text data need to be added. Only four automobile makers' quarterly call transcripts from 2014-2019 were collected. Due to the small data set, this model is subjected to a limited use in a few automobile manufactures.
2. Neural network does not provide value to improve accuracy; also, lack of interpretation may make neural network not suitable for majority of investors.
3. Try to make more classes: ideal candidates could be market is up, down, flat. Alternative classification could be negative or positive value with a discrete range of 0-10%, 10-20%, 30-40%, 50% or above. After all, survey what is the probability of each prediction.