

Ada Kilic

Report 3 (deadline: Jul 3, 23:59):

- Description of the measurement strategies;

In the context of my research, measurement strategy means how I plan to quantify sentiment and subsequently the overall sentiment index from the text data of Chair's statements from the FOMC meetings. Since I am using FinBERT which is a model specifically used for financial sentiment analysis, the measurement strategy is largely dependent on the outputs that this model provides.

Firstly, sentence-level sentiment analysis will be measured. This approach will analyze the sentiment of each sentence in the policy statements. This involves applying FinBERT to each sentence individually and obtaining a sentiment score. These scores can then be averaged to obtain an overall sentiment score for the statement. Also another approach is to analyze the sentiment of the entire policy statement as one document. This could potentially capture more holistic sentiment but will miss out on many factors therefore I will not follow this approach. But it could be useful to use as a comparison. Additionally, a measurement strategy is weighted sentiment scores. This where some sentences or sections of a statement are more impactful or important than others. Therefore, I may consider a weighted average sentiment score. For instance, the summary or conclusion sentences might carry more weight. Also depending on the sentiment scores' distribution, I will apply a threshold to classify the sentiment into categories such as positive, negative, and neutral which can simplify the interpretation of the results. I would also like to measure the change in sentiment from one statement to the next, instead of absolute sentiment scores. This could help identify trends or significant shifts in sentiment over the years. Lastly, another measurement strategy is exploring sentiment dispersion. This is an interesting measure that could be the variance of sentiment scores within a statement. A high dispersion might indicate a statement that discusses both positive and negative aspects.

- Statistical Analysis (similar to our in-class work);

Statistical analysis is crucial to establish relationships, trends, and identify significance in my text dataset. For my I am constructing a sentiment index from policy statements and trying to understand its signal on the financial markets. I will use correlation analysis. Correlation analysis can guide me to identify whether there's a linear relationship between the sentiment index and key financial market indicators. There are a few different correlation analyses like Pearson correlation that can be used. I may also do a regression analysis to understand the impact of the sentiment index on financial markets while controlling for other variables. Furthermore, Time Series Analysis can be useful for my project. Given that my data will increase over the years, statistical techniques specific to time series data can be used. Also change point analysis can be used. This is a technique that can help detect when a significant change has occurred. For instance, it can be important to identify points in time where the sentiment dramatically shifts (eg.

maybe this happens during a recession) and then analyzing what was happening in financial markets at that time. I will also do a hypothesis test for instance a positive sentiment is associated with an increase in the stock market. Lastly, examining how much sentiment fluctuates could give interesting insights. Variance or standard deviation can be used to measure this. Additionally, the relationship between sentiment volatility and financial market volatility can be explored.

- Visualization of the data. You need to describe your number visually and have one separate page for the visual storytelling.

Some data visualizations I am considering includes:

Time Series Plot of Sentiment Index: This can be used because I am creating a historical sentiment index. This can show how sentiment changes over time. The x-axis can represent time (from 2000 to present), and the y-axis can represent the sentiment index. Each point on the graph represents the average sentiment value of each FOMC statement.

Heatmap of Sentiment over Time: A heatmap can help me show the sentiment of each FOMC meeting over the course of a year. The x-axis could represent months, the y-axis could represent years from 2000 to present. Finally, the color could represent sentiment with a gradient from negative to positive.

Scatterplot: A scatterplot of sentiment index vs. a financial variable can help visually analyze the relationships.

Word Cloud: Although this isn't strictly a visualization of the sentiment scores, a word cloud of the most common words in positive and negative sentences could provide some insights into what kinds of topics or phrases are associated with different sentiments.

Change Point Analysis Visualization: If I conduct a change point analysis, it would be useful to overlay the identified change points on a time series plot of the sentiment index.