

Implication Assessment Model Using Zero-Shot Text Classification to Interpret Fed Monetary Policy Communication

Analysis of the Fed' s communication by using textual entailment
model of Zero-Shot classification

Satoshi Sawaki¹ Yasuhiro Nakayama²

Tomochika Sawaki¹, Yasuhiro Nakayama²

¹Mizuho Bank

¹Mizuho Bank, Ltd.

²Mizuho Research & Technologies

² Mizuho Research & Technologies, Ltd.

Abstract: This study uses text mining techniques to analyze documents published by central banks and propose a method for evaluating the policy tone of central banks. Since it also has a significant impact on the real economy, market participants are trying to better grasp changes in the central bank's future monetary policy outlook. Sentiment analysis of central bank documents has been conducted for a long time, but it has been difficult to accurately interpret the meaning of the text and to explicitly capture intentional changes in nuance. In this study, by using the implication determination method based on zero-shot text classification, we intend to continuously evaluate the unknown economic environment using the same model, and the statements and minutes of the FOMC meeting released by the Fed, Tone changes were compared when each of the press conference transcripts and Fed high-ranking remarks were used as data sources. In addition,

Minutes of FOMC meetings This is a phase analysis of changes in each policy stance since 1971.

1 Getting started

Since the monetary policy of major central banks has a wide range of influences, such as financial market developments, pricing of risk assets, and spillover to the real economy, market participants are trying to better grasp changes in the central bank's future monetary policy outlook. The monetary policy of the central bank of the United States (Federal Reserve, hereinafter referred to as the Fed) is positioned as the most important because it affects the movement of the dollar, which is the key currency. One of the means for central

banks to engage in dialogue with the market and conduct policy smoothly is the publication of various documents, such as statements and minutes released after the policy meetings, and transcripts of speeches and congressional testimonies attended by senior officials. . The Federal Open Market Committee (FOMC), the meeting where U.S. monetary policy is formulated, consists of Federal Reserve (Fed) members and regional Fed presidents. It was held eight times, and immediately after the meeting, a statement was published on the website, and a press conference was held by the chairperson, and a transcript of the press conference

(Transcript) is published. Minutes of FOMC meetings will be published about three weeks later (until December 2004, 3 days later). The statement is a relatively short document of about two pages that summarizes the current economic perception, a summary of the monetary policy decided, and the names of the voters. In addition to the transcript read by the chairperson at the beginning, the transcript of the press conference consists of a question-and-answer session with reporters, and the content is 20~30. It will be before and after the page. In some cases, content that is not stated in the statement but is of interest to market participants (specific content and future prospects) may be recorded. Following the report of the economic analysis by the Fed economist, the minutes are documents that confirm the discussion process that led to the decision on policy and the differences of opinion among the members. The content is about 10~20 pages. Outside of FOMC meetings, transcripts of speeches, interviews, and congressional testimonies by FOMC members (Fed officials) are published each time they are conducted. Although it may not necessarily be related to monetary policy, if there is a significant shift from past statements by senior Fed officials to express their own perceptions of the economy and monetary policy outlook, it may have an impact on the market. Statements by Fed members are published on the website, and statements by regional Fed presidents are published on the websites of each Fed. The content is often around a few pages. Another item that attracts the attention of market participants is the Beige Book (District Fed Business Report). This is a report compiled by the Regional Fed on the local economy of all 12 districts. The report was released on Wednesday, two weeks before the FOMC meeting, and will be used as a reference at the FOMC meeting. substance

The amount will be around 30 pages.

The format of each type of document is characterized by the fact that statements, minutes, and beige books are documents in which the structure of contents and chapters are stylized, while conference transcripts and statements by high-ranking officials are non-standardized documents. In terms of implications for monetary policy, it is generally considered that the manuscript of high-level statements is ahead of the curve in that it allows us to grasp the perceptions of senior officials in advance of the FOMC meeting (however, the FOMC There is a blackout period immediately before the meeting, and no high-ranking statements are made). In addition, it is considered that more information can be obtained in meeting minutes and press conference transcripts with a large amount of text compared to statements with a small amount of content.

Therefore, in this study, we conducted text mining using statements, meeting minutes, conference transcripts, and statements by high-ranking officials, and compared their usefulness as information sources.

2 Related research

As an application of text mining technology to the financial and economic fields, many studies of central bank documents have been conducted in the past, as well as analysis of corporate financial results and new accounts. Many studies have used text mining techniques for central bank documents to score the central bank's sentiment and to predict future monetary policy and market forecasts. Many of them are active for the purpose of forecasting economic indicators.

Ito et al. combined FOMC-specific dictionaries prepared by experts with engagement analysis for FOMC minutes to extract sentiment by topic. It was shown that it has explanatory power for macroeconomic indicators[1]. Wang uses sentence-level embedding vectors obtained from FinBERT to calculate topical sentiment. They

were shown to have explanatory power over macroeconomic indicators[2].

Granziera et al. also calculated sentiment toward inflation using speech texts of FOMC members and district Fed chairmen . [3] There is also a study that analyzes the sentiment of each member for transcripts that record all participants' remarks, published five years after the FOMC meeting. [4].

The purpose of this study is to use natural language processing technology to read the intentions of document writers and speakers as correctly as possible, and to become the most correct recipient of communication intended by FED. Therefore, since 1993, when the minutes were first published in their current format, and the past minutes that had been published in a similar format can be traced back to 1971 . We analyzed the characteristics of the policy tone of the fed over a long-term span by separating it according to the trend of the policy interest rate.

3 Proposed method

3.1 Data acquisition and preprocessing

FOMC ¹statements, minutes, press conference transcripts, and statements by senior officials published on the Fed's website were obtained. Paragraphs that are not related to stance, such as the names of participants and explanations of specific details of formulated policies, were excluded from the analysis. (However, all paragraphs are covered for older minutes for which the paragraph structure cannot be obtained.) Although only the transcript read aloud by the chairperson on the day of the press conference is released and the next day it is released including the question-and-answer portion by the reporter, this time the transcript including the

question-and-answer part was analyzed. . However, for the question-and-answer part, the reporter's question text was deleted and only the chairman's response was extracted.

The retrieved documents were divided into sentence units after preprocessing such as removing hoods and annotations.

3.2 Topic determination by implication determination model

Topic judgment is performed using the zero-shot base [6] implication judgment model. In this study, we used a trained model that is publicly available. When selecting models, members with domain knowledge compared the outputs of several models and selected the one that best suited their senses². The implication determination model returns a score of 0~1 to determine whether the sentence to be judged implies the meaning of the hypothetical statement. A score closer to 1 means that it implies a hypothesis statement, and a score closer to 0 means that it contradicts the hypothesis statement. Since it is a zero-shot model, hypothetical statements can be set freely. This time, we will call the hypothetical statement "This sentence is related to the topic of {}." and set the topic you want to judge in parentheses. Topic classification was performed.

The topics of this time are "Inflation", "Employment", "Economic Activity)" and judged each sentence. Since multiple topics may be included in a sentence, it is judged whether or not each topic belongs to it. Since the model returns an implicit judgment score, a threshold value is set, and if the score is above the threshold, it is judged that the topic belongs to the topic. The threshold was set to 0.9 by members with

¹ <https://www.federalreserve.gov/default.htm>

² <https://huggingface.co/facebook/bart-large-mnli>

domain knowledge who confirmed the judgment scores of several sentences extracted as a sample.

3.3 Setting Hypothesis Statements by Category

In order to determine the implication of detailed nuances for each category, a hypothesis statement for each category is created. The hypothesis statement was set by the following process.

First, from the previous statements published by the Fed, "Inflation", "Job Gain", "Economic Growth" Expressions for three categories were extracted on a keyword basis. Frequently used expressions were selected from the extracted expressions for each category and set as hypothetical statements. When choosing an expression, the balance of direction was also considered. For example, about "Inflation" "declined", "diminished", "edged down" as the downward direction, "increased", "moved up", "elevated", etc. Expressions in both directions are extracted in a well-balanced manner.

3.4 Implication judgment by category

Based on the set categorical hypothesis, the implication judgment is made for each category. Based on the results of the topic judgment in 3.2, the target sentences for the connotation judgment for each category were as follows.

"Inflation" Category For "Inflation" (Inflation) "Topics, "Job Gain" Category "Employment" Topics, "Economic Growth" Categories were analyzed for sentences judged to belong to the "Economic Activity" topic. The model for determining connotations for each category was the same as that of topic judgment in 3.2. Implication judgment is performed on the categorical hypothesis sentence set in 3.3, and the connotation judgment score is the threshold Count the number of sentences that were greater than or equal to (0.9) and aggregate them for each document. In each

categori, the number of implied sentences in each direction of expression is counted, and the difference is taken and divided by the total number of implicit sentences in both directions to achieve a stance score $S_c(d)$.

$$S_c(d) = \left(\sum_{e \in E_{c,p}} C_e(d) - \sum_{e \in E_{c,n}} C_e(d) \right) / \sum_{e \in E_{c,p}, E_{c,n}} C_e(d)$$

where $E_{c,p}$ is the set of upward direction representations in category c , $E_{c,n}$ is the set of downward direction representations in category c , $C_e(d)$ is represented in document d It is the number of sentences judged to imply e .

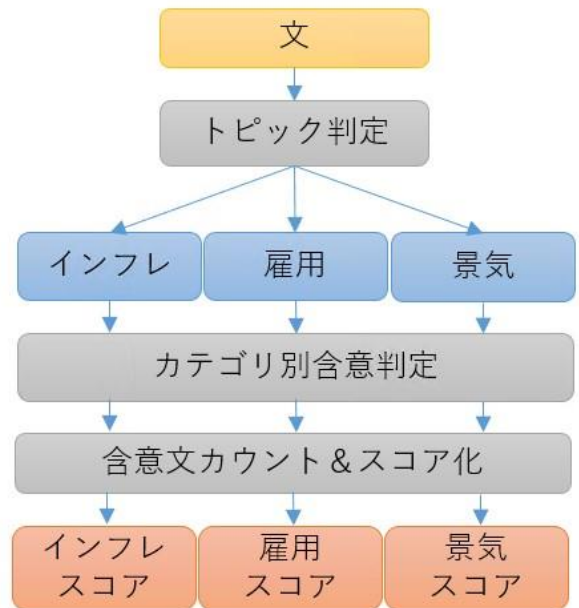


Figure 1: Analysis Flow

4 Results and discussion

4.1 Data

From the Fed website, FOMC statements, minutes, conference transcripts, and high-level remarks were obtained, and evaluations were made. Minutes taken from February 1971 only, others from December 2018. In addition, since it is not possible to verify the changes in policy after February 2023,

which is the time of this study, March 2022. The analysis was performed separately up to the month.

4.2 December 2018 to March 2022 analysis

FOMC statements, minutes, and transcripts of FOMC meetings from December 2018 to March 2022. Using data from senior officials' remarks, the phase after 2021 (amid rising inflation, the Fed has shifted from an accommodative policy of zero interest rate policy and quantitative easing to tapering and raising policy rates, B/S We conducted text mining of the phase of shifting to a tightening policy such as contraction) and compared each data source. For comparison, the comparison was limited to December 2018 when the chair held a press conference after each FOMC meeting. This is to match the timing.

Figure 2 plots stance scores on inflation. Hypothetical representation of downward inflation from the percentage of sentences judged to imply a hypothesis of an upward trend in inflation (e.g., "increased", "picked up", "moved up", "elevated"). ("declined", "diminished", "moved lower", "edged down" etc.).

Regarding the FOMC statement, it was indicated that the inflation rate was on a downtrend until the March 2021 meeting. The score seemed to indicate that upward pressure on inflation was increasing since the monthly meeting. Then, a similar trend followed.

Next, regarding the score using FOMC minutes as a data source, the score moved from the perception of low inflation to near neutral from the September 2020 meeting. The recognition that upward pressure on inflation was intensifying began at the April 2021 meeting, as in the statement. Then, a similar trend followed. However, it should be noted that the

minutes are published three weeks after the meeting.

Ru.

Third, looking at the changes in scores using FOMC conference transcripts as a data source, as in the case of FOMC meeting minutes, scores are gradually returning to near neutral since the September 2020 meeting. At the March 2021 meeting, the Bank came to recognize that upward pressure on the inflation rate was increasing. This was the earliest turning point than the statements and minutes. At the March 2021 meeting, although there was no significant change in the recognition of inflation in the statement, the Chairman's press conference after the meeting. This result suggests that the perception of increasing inflationary pressure may have been demonstrated. Then a similar trend continued. Given that the minutes are published three weeks after the meeting, while the transcripts of the press conferences are released the next day, the fastest of the three data sources.

It is thought to have been exchanged.

Finally, it is about the case of using high-ranking officials' statements as sources. For comparison with other documents, the average of the five statements made by senior officials just before the FOMC meeting is provided. As with the minutes, the score gradually started to increase from September 2020 and increased sharply in April 2021.

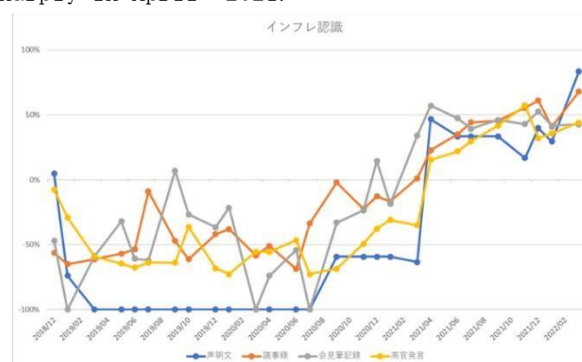


Figure 2: Stance on Inflation

Next, we check the results of the similar analysis of employment shown in Figure 3. With regard to the transition to monetary easing following the shutdown in response to the global spread of the novel coronavirus since March 2020, Due to the emergency rate cut, all sources confirmed a sharp decline in scores in April 2020, and the perception that the job market is accommodative has shifted. Since then, the transcripts and minutes of the press conference have gradually returned to neutrality, while the statements have been published in April 2021. At the March meeting, the employment environment shifted to a tight perception with a near-digital change range.



Figure 3: Changes in Stance on Employment

Finally, the economic perception is shown in Figure 4. In the February 2019 meeting statement and minutes, the economic perception turned to a slowdown, July 2019. It was confirmed that the policy interest rate cut was preceded by the rate cut after the month. Around March 2020, as with employment perceptions, the economic slowdown after the shutdown. There was no significant difference in the timing of the score decline due to the difference in data sources. As for subsequent developments, scores created based on meeting minutes, transcript records, and statements by high-ranking officials were published around

the summer of 2020, 2020. While the score has returned to the same level as before March, the statement. It returned to neutral scores in spring 2021.



Figure 4: Stance on the Economy

4.3 Analysis using minutes from February 1993

Table 1 shows the minutes of FOMC meetings from February 1993 to March 2022 as data sources. Scores were generated in the same way as in 4.2 and compared for each aspect of policy stance. Federal Fund Rate is available on the Fed website.

and obtained from Bloomberg.

First, in terms of inflation perceptions, since 1993, the inflation rate (annual core PCE) will start at 1% through 2021. In a calm situation at a low level that hovers between 3 percent, the communication of the Fed was also recognized as having low upward pressure on inflation on average for all periods. On the other hand, the score was higher than the average for all periods during the period when the policy interest rate was raised, and lower than the average for the entire period during the period when the interest rate was cut and the period when the interest rate was zero. This result is consistent with the idea that the policy interest rate is changed as a means to achieve the inflation target or stabilize prices.

Next, the employment recognition is shown in Table 1. The score was higher than the average for all periods during the period when the policy

rate was raised, and lower than the average for all periods during periods of rate cuts and zero interest rates. This result is consistent with the idea that the policy rate is changed as a means to achieve employment maximization, and it is considered that the communication of the Fed can be read correctly. .

Finally, the same trend was confirmed for the score related to business perception. When viewed as an average for all periods, the score was positive, that is, it was recognized that the economy was strong, and the score was higher than the average for the entire period during the interest rate hike phase. On the other hand, during rate cuts, the score is negative on average, and when the Fed perceives that the economy is in a recession. It was suggested that the time at which the policy interest rate was actually lowered generally coincided.

Table 1: Period Average of Scores Using FOMC Minutes

	①	②	③	④
inflation	-0.23	+0.00	- 0.40	-0.34
employment	-0.08	+0.26	- 0.45	-0.24
business	+0.22	+0.44	- 0.16	+0.20

*① Entire period, ② Interest rate hike period, ③ Interest rate cut period, ④ Zero interest rate period

(2009/1~2015/10, 2020/4~2022/1)

*2 Interest rate hikes are included in the interest rate hike period even for meetings where interest rates are left unchanged. * Welch's alternative hypothesis is that "(2) the average value of the rate hike phase > (3) the average value of the rate cut phase" is the alternative hypothesis of Welch's t. In the test, inflation, employment, and economy are all

The significance level of 1% was judged to be a significant difference.

4.4 Analysis using minutes from

February 1971 to December 1992

Before December 1992, there were no minutes in their current format, but there was a document similar to the minutes of action. Table 2 shows the Minutes of FOMC meetings from February 1971 to December 1992. Using Action as a data source, comparisons are made in the same manner as in 4.3 for each aspect of policy stance. Although the score level was different, the same trend as 4.3 was confirmed. Comparing the all-time averages of Tables 1 and 2, Table 2 has a higher score for inflation perception, while a lower score for employment perception is higher. The recognition scores matched. This is consistent with the relationship between the average inflation rate (core PCE) and the unemployment rate for each period. In addition, Table 2 shows that the average values of inflation, employment, and the economy tend to be higher in the rate hike phase than in the rate cut phase, as in Table 1. Thus, over the past 50 years, the Fed's response stance to inflation and employment data has been consistent. These results suggest that it may be quantified using the model in this study.

Table 2: Period-to-period average of scores using FOMC minutes

	①	②	③
inflation	-0.05	+0.17	-0.34
employment	-0.31	-0.13	-0.63

business	+0.2 2	+0.3 0	+0.0 8
----------	-----------	-----------	-----------

*(1) Entire period, (2) Interest rate hike period, (3) Interest rate cut period

* In Welch's t-test, which was based on the alternative hypothesis of "(2) the average value of the rate hike phase > (3) the average value of the rate cut phase" showed that inflation, employment, All economies were judged to have a significant difference at the significance level of 1%.

5 Conclusion

In this study, we attempted to read changes in Fed communication using an implicit judgment model based on zero-shot text classification. Since zero-shot classification is used, it is possible to respond to other themes without additional learning.

Using data from December 2018 to March 2022, we attempted to compare each document published by the Fed. Regarding statements, the scores generated by the model used in this study change scores with somewhat digital movements, while for meeting minutes and transcripts, A more gradual change in stance was confirmed. In addition, the average of the past five statements by high-ranking officials has been taken, and they have been moving around the neutral level. As a future issue, it is possible to devise data preprocessing.

Next, we analyzed the phases of the FOMC meeting using the minutes of the FOMC meeting for the long-term period from February 1971 to March 2022, and found that inflation and employment Using the model in this study, it was possible to interpret that the Fed communication on business perception is consistent with actual policy changes. .

thanks

This research was inspired by the first half report of the Data Science school of the University of Tokyo. I would like to thank all the students who worked diligently on the assignments and came up with wonderful ideas, as well as the TAs and supervisors who guided me.

Considerations

The contents and views of this paper belong to the individual author and affiliation. It is not the official position of the company.

bibliography

- [1] Ryo Ito, Shintaro Suda, Kiyoshi Izumi: Evaluation of FOMC Minutes by Topical Polarity Method , SIG-FIN-017, (2016)
- [2] Sarah-Yifei-Wang: Aspect-based Sentiment Analysis in Document - FOMC Meeting Minutes on Economic Projection, arXiv:2108.04080, (2021)
- [3] E Granziera, V. H. Larsen, G Meggiorini: Fed Sentiment and Expectations: Evidence from Speeches by FOMC Members, (2022)
- [4] S Cannon, Sentiment of the FOMC: Unscripted, economic Revie, 2015, issue Q IV, 5-31, (2015)
- [5] L Xiao: Using Sentiment Analysis to Understand Monetary Policy Uncertainty, (2022)
- [6] W Yin, J Hay, D Roth: Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach, In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP/IJCNLP 2019, Hong Kong, China, November 3-7, 2019, pp. 3914-3923, (2019)