



Project 1: Deciphering FedSpeak with NLP

Background

The [Federal Reserve \(the Fed\)](#) is the central bank of the United States. Its role is to conduct monetary policy to promote the health of the US economy and the stability of the US financial system. Within the Fed, the [Federal Open Market Committee \(FOMC\)](#) meets eight times a year to discuss the economic outlook and set the direction for monetary policy. These meetings are followed by public statements that summarise the committee's view of the economy and monetary policy decisions.

FedSpeak is a term used to describe the language used by the Fed to communicate on monetary policy decisions and manage investors' expectations about future policy directions. The sentiment in the Fed's communication can be classified as hawkish or dovish, where hawkish means the Fed is more likely to tighten monetary policy/raise interest rates to curb inflation and cool down the economy, while dovish means the Fed is more likely to loosen monetary policy/cut interest rates to support faltering growth. Depending on how hawkish or how dovish it is, the sentiment in the Fed's language could significantly swing financial markets. As a result, central bank statements, meeting minutes and individual member speeches have become important data to be analysed by investment professionals.

Besides the Fed, other major central banks that exert great market influences include the [European Central Bank \(ECB\)](#), the [Bank of England \(BoE\)](#), the [Bank of Japan \(BoJ\)](#) and so on. The sentiment in their communications is important to market watchers too.

Problem

Most of the time deciphering central bank communications is more like an art than science, with interpretations subject to human biases. For optimal decision making, it is essential to develop a way to systematically extract information from central bank communications and objectively quantify any policy signal. This could be possible with the help of Natural Language Processing (NLP), especially with the rise of Generative Pre-trained Transformer (GPT) models, which already have some relevant domain knowledge. As shown in Hansen & Kazinnik (2023), the performance of GPT models in classifying the Fed's policy stance surpasses that of other popular classification methods.

However, without further fine-tuning, the performance of GPT-3 Zero-Shot Learning is still not good enough to match human analysts. How to most cost- and time-efficiently utilise the pre-trained base models to improve performance in central bank communication sentiment classification is a challenge for Andromeda.

Dataset

The provided dataset includes two files:

1. FOMC Statements 1997-2023.xlsx: This file contains the text of all FOMC statements released after meetings since 1997.
2. FOMC Labelled Sentences.xlsx: This file contains 200 randomly-drawn sentences from all the statements. Each sentence has been pre-labelled by human analysts with a



hawkishness/dovishness score as defined in the table below. There are also some sentences labelled as “Remove” which indicates sentences that are irrelevant to monetary policies and should be removed in the data cleaning stage.

Sentiment	Score	Definition
Dovish	-1	Strongly expresses a belief that the economy may be growing too slowly and may need stimulus through monetary policy
Mostly Dovish	-0.5	Overall message expresses a belief that the economy may be growing too slowly and may need stimulus through monetary policy
Neutral	0	Expresses neither a hawkish nor dovish view and is mostly objective
Mostly Hawkish	0.5	Overall message expresses a belief that the economy is growing too quickly and may need to be slowed down through monetary policy.
Hawkish	1	Strongly expresses a belief that the economy is growing too quickly and may need to be slowed down through monetary policy

Besides the provided data, you are also encouraged to collect other central bank communication data to supplement the study. Examples include FOMC meeting press conference transcripts, FOMC meeting minutes, individual member speeches or statements from other central banks. These could be found on the central banks’ websites:

- FOMC press conference transcripts and meeting minutes: <https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm>
- FOMC member speeches: <https://www.federalreserve.gov/newsevents/speeches.htm>
- ECB statements: <https://www.ecb.europa.eu/press/pressconf/html/index.en.html>

Project Tasks/Deliverables

You are invited to help Andromeda design a system that functions as a junior research analyst who can quickly process central bank communications and provide a hawkishness/dovishness assessment. You are free to choose your training methods or the ways you utilise GPT base models, subject to the budget cap provided (US\$ 300). However, efficiency is as important as accuracy - you need to consider the ongoing costs and efforts for Andromeda in using the system. In designing the overall system, the following two steps must be included:

1. Evaluate the performance of different classification methods on the pre-labelled dataset to choose the best one.
2. You need to aggregate the scores generated using the most optimal classification method for all sentences in each statement to generate an overall hawkishness/dovishness score for each Fed meeting and construct a time series of Fed sentiment indicator for all meetings in 1997-2023.

You should keep a clear record of the different methods you used, the respective training processes and your codes.

Additional Reading

- Hansen, A. and Kazinnik, S., [Can ChatGPT Decipher Fedspeak?](#), March 2023
- Pan, T., and Lee, H., [AI in Finance: Deciphering Fedspeak with Natural Language Processing](#), March 2021