Chapter 1

Project and Goal

1.1 Executive Summary

The primary objective of this project is to analyze the impact of the American central bank's (Federal Reserve) speeches on market price movements, specifically focusing on the S&P 500 index. This analysis will employ a Natural Language Processing (NLP) artificial intelligence model to examine and extract sentiment from statements made by Federal Reserve representatives. By detecting underlying sentiment, we aim to explore any correlation between the tone of these speeches and subsequent price movements in the market.

Given the complexity and often ambiguous nature of central bank communications, this project seeks to determine if sentiment extraction from these speeches can reveal a clear link to market responses. This analysis will help to clarify whether market movements are aligned with the sentiment conveyed or if they exhibit a more random pattern.

1.2 Output

The expected outcome of this project is a comprehensive report that summarizes and documents the entire process, from dataset creation and sentiment analysis to the final interpretation of the results. This report will detail the methodology for data collection and processing, sentiment analysis approach, and findings on the correlation, if any, between extracted sentiment and market movements. The report will also address whether the observed market reactions are consistent with the derived sentiment or appear independent of it.

1.3 Motivations

In financial markets, central bank communications are pivotal, often resulting in heightened speculation and volatility. However, the complex language used in these communications can lead to significant uncertainty and varied interpretations among market participants. By analyzing the sentiment within the Federal Reserve's speeches, we aim to uncover whether these statements hold predictive power over market movements or if they merely introduce noise.

This project also serves as an opportunity to advance our understanding of Natural Language Processing in financial contexts, focusing on sentiment analysis as a tool for market insight.

Chapter 2

Tasks and Resources

2.1 Method Identification

The project consists of several phases: gathering speech transcripts, pre-processing text data, applying NLP models to extract sentiment, and analyzing the relationship between sentiment scores and S&P 500 price changes.

2.2 Prerequisites

A strong understanding of Natural Language Processing techniques, financial market dynamics, and sentiment analysis methodologies is essential. Proficiency in Python and familiarity with libraries such as Pandas, NLTK for NLP.

2.3 Frameworks, APIs, and Languages

- **Programming Languages:** Python
- **Libraries:** Pandas, NLTK, Matplotlib, Seaborn
- **Source Code Editor:** VSCode
- **Data Source:** Scarping, Federal Reserve Speech Transcripts
- **Documentation Platform:** Overleaf, Zotero

Chapter 3

Timeline

The Roadmap as gaant is coming soon...

3.1 Roadmap

The project roadmap is divided into several weekly phases. Key tasks include data collection, model selection and training, sentiment extraction, and data visualization. Weekly meetings will track progress, discuss findings, and address challenges.