Overview

1. Central bank speech archive.
   ▶ Maintained by Bank for International Settlements (BIS).

2. Methods and models for text feature extraction.
   ▶ Collection of large language models (LLMs).

3. Description and visualization of text features from speeches.
Textual Data
Textual Data

Overview

- **BIS speech archive.**
  - Central bank speeches translated into English.
  - >18,000 speeches; >90 institutions.
  - Covers late 1996 to present.

- **Our coverage.**
  - 53 institutions with >50 speeches.
  - Extract text features using LLMs at paragraph level.
  - Aggregate to quarterly or annual features.
Textual Data

Database Coverage

International Central Bank Communication
Methods and Models
Methods and Models

Reference

Methods and Models

Text Feature Extraction: LLMs and Methods

1. **Transformer model.**
   - Maps sequence of embeddings to sequence of contextualized embeddings.

2. **Sentence transformer model.**
   - Maps sequence of embeddings to single embedding.

3. **Feature extraction methods.**
   - Zero shot classification (ZSC), extractive question answering (EQA), and semantic textual similarity (STS).
Zero Shot Classification

**Sequence:** “Banks continue to play this role but it has become more challenging today to do so because some lenders find themselves capital constrained as a result of recent losses and or sizable unanticipated additions to their balance sheets of formerly off balance sheet instruments.”

**Candidate Classes:** ['financial stability', 'output', 'inflation', 'labor market']

**Scores:** [0.718, 0.203, 0.048, 0.031]
# Zero Shot Classification: Attention Mechanism Performance

<table>
<thead>
<tr>
<th>[CLS]</th>
<th>[CLS]</th>
</tr>
</thead>
<tbody>
<tr>
<td>commercial</td>
<td>commercial</td>
</tr>
<tr>
<td>banks</td>
<td>banks</td>
</tr>
<tr>
<td>have</td>
<td>have</td>
</tr>
<tr>
<td>long</td>
<td>long</td>
</tr>
<tr>
<td>been</td>
<td>been</td>
</tr>
<tr>
<td>thought</td>
<td>thought</td>
</tr>
<tr>
<td>of</td>
<td>of</td>
</tr>
<tr>
<td>,</td>
<td>,</td>
</tr>
<tr>
<td>and</td>
<td>and</td>
</tr>
<tr>
<td>indeed</td>
<td>indeed</td>
</tr>
<tr>
<td>have</td>
<td>have</td>
</tr>
<tr>
<td>functioned</td>
<td>functioned</td>
</tr>
<tr>
<td>as</td>
<td>as</td>
</tr>
<tr>
<td>,</td>
<td>,</td>
</tr>
<tr>
<td>he</td>
<td>he</td>
</tr>
<tr>
<td>backup</td>
<td>backup</td>
</tr>
<tr>
<td>source</td>
<td>source</td>
</tr>
<tr>
<td>of</td>
<td>of</td>
</tr>
<tr>
<td>liquid</td>
<td>liquid</td>
</tr>
<tr>
<td>#ity</td>
<td>#ity</td>
</tr>
<tr>
<td>for</td>
<td>for</td>
</tr>
<tr>
<td>many</td>
<td>many</td>
</tr>
<tr>
<td>other</td>
<td>other</td>
</tr>
<tr>
<td>financial</td>
<td>financial</td>
</tr>
<tr>
<td>institutions</td>
<td>institutions</td>
</tr>
<tr>
<td>and</td>
<td>and</td>
</tr>
<tr>
<td>markets</td>
<td>markets</td>
</tr>
</tbody>
</table>
Extractive Question Answering

**Query 1:** What is the most significant concern in the passage?

**Context 1:** “The suspension of the convertibility of the dollar into gold and the imposition of a 10 percent import surcharge last summer ran the risk of mass foreign retaliation in the form of destructive trade barriers.”

**Output 1:** mass foreign retaliation
Methods and Models

Semantic Textual Similarity

- Use contextualized sentence embeddings and semantic textual similarity.
  - Train using Siamese and triplet networks (Schroff et al., 2015).

- Compare sequences from speeches with descriptions of policy objectives or preferences.
  - “Monetary policy should be used to achieve financial stability.”
  - “Banking regulation should be used to achieve financial stability.”
Methods and Models

Extended Pretraining with TSDAE

Figure 1: Architecture of TSDAE.

Figure taken from Wang et al. (2021).
Refine STS Performance with Fine-Tuning

1. Use S2ORC abstracts (Lo et al. 2020).

2. Randomly draw similar sequence pairs from the same paper abstract.

3. Randomly draw dissimilar sequence pairs from different abstracts.

4. Train on STS and compare using cosine similarity.

\[
    \text{sim}(S_i, S_j) = \frac{S_i \cdot S_j}{\|S_i\| \|S_j\|}
\] (1)
Methods and Models

Non-Dual Mandate Content: 1984-2020

Methods: Zero shot classification and extractive question answering.
Methods and Models

Cosine similarity: Banking Regulation and Financial Stability

Methods: Zero shot classification and semantic textual similarity.
Methods and Models

t-SNE Plot: Speech Embeddings for Financial Stability Content

Methods: Zero shot classification and semantic textual similarity.
Dataset

Textual Features

► **Text classification.**
  
  ► Financial stability, output and employment, inflation, and exchange rate.

► **Semantic textual similarity.**

  1. **Individual features:** Monetary policy, financial crisis, bank regulation, bank capital and liquidity, U.S. dollar, international trade, hawkish sentiment.

  2. **Policy advocacy:** monetary policy and financial stability, bank regulation and financial stability.
## Dataset

### Documentation

<table>
<thead>
<tr>
<th>Sheet Name</th>
<th>Coverage</th>
<th>Description</th>
<th>Method</th>
<th>Transformation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>bankreg_finstab</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Semantic textual similarity.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>monetary_finstab</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Semantic textual similarity.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>financial_crisis</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Semantic textual similarity.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>united_states_dollar</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Semantic textual similarity.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>bank_regulation</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Semantic textual similarity.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>monetary_policy</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Semantic textual similarity.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>hawkish_sentiment</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Semantic textual similarity.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>bank_concern</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Semantic textual similarity.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>international_trade</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Semantic textual similarity.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>output_employment</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Zero shot classification.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>inflation</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Zero shot classification.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
<tr>
<td>exchange_rate</td>
<td>1996-2023</td>
<td>This feature measures the extent to w Zero shot classification.</td>
<td>Standardized using mean and standard deviation for 1996-2010 period.</td>
<td>Bertsch, Hull</td>
<td></td>
</tr>
</tbody>
</table>
Documentation

- Text features standardized using mean and standard deviation calculated over 1996-2010 period.
  - Updates to series will not affect past values.

- Central banks split into annual and quarterly groups.
  - Quarterly features provided for institutions that give at least two speeches per quarter on average.

- No data imputation.
  - Many central banks lack sufficient number of speeches over 1996-2023 period to yield text features for all periods.
Dataset

Heatmap of Exchange Rate Discussion

International Central Bank Communication
Inflation Feature Principal Components

Date
Rolling Component Value
PC1
PC2

International Central Bank Communication
Inflation Feature Correlation with PC1

- Swiss National Bank
- Reserve Bank of India
- Bank of Thailand
- Bank of Albania
- Central Bank of Malaysia
- Bank of Japan
- Bank of France
- Norge’s Bank
- Bank of Canada
- Central Bank of Ireland
- Bank of England
- Deutsche Bundesbank
- Bank of Spain
- Monetary Authority of Singapore
- Sveriges Riksbank
- Bank of Italy
- South African Reserve Bank
- US Federal Reserve System
- European Central Bank
- Bank of the Philippines
- Reserve Bank of Australia

Correlation Coefficient

0.0

Swiss National Bank
Reserve Bank of India
Bank of Thailand
Bank of Albania
Central Bank of Malaysia
Bank of Japan
Bank of France
Norge’s Bank
Bank of Canada
Central Bank of Ireland
Bank of England
Deutsche Bundesbank
Bank of Spain
Monetary Authority of Singapore
Sveriges Riksbank
Bank of Italy
South African Reserve Bank
US Federal Reserve System
European Central Bank
Bank of the Philippines
Reserve Bank of Australia
Discussion of Bank Regulation
US Federal Reserve System
European Central Bank
Reserve Bank of India
Bank of England
Bank of Japan
Next Steps

Discussion

➤ Models and methods can be applied broadly to extract features from central bank texts.
  ➤ Text classification, extractive question answering, semantic textual similarity.

➤ Additional policy advocacy features under development and available on request.
  ➤ Capital controls and exchange rates, capital controls and financial stability, monetary policy and exchange rate.

➤ Open to accepting input for features to extract from texts.
  ➤ Can make features available in updates.