**IFC-Bank of Italy Workshop on "Data science in central banking: enhancing the access to and sharing of data"**

**17-19 October 2023**

# Let’s talk about sentiment: natural language processing using machine learning on bank earnings transcripts[1]

Seung Jung Lee, Sriram Nagaraj, Dylan Saez,
Victors Stebunovs, Cindy Vojtech and Karl Wirth,
Board of Governors of the Federal Reserve System

[1] This contribution was prepared for the workshop. The views expressed are those of the authors and do not necessarily reflect the views of the Bank of Italy, the BIS, the IFC or the other central banks and institutions represented at the event.

# Let's Talk About Sentiment: Natural Language Processing Using Machine Learning on Bank Earnings Transcripts

**Seung Jung Lee**, Sriram Nagaraj, **Dylan Saez**,
Victors Stebunovs, Cindy Vojtech, and Karl Wirth

3rd IFC and Bank of Italy Workshop on
"Data Science in Central Banking: Enhancing the access to and sharing of data"

19 October 2023

The views expressed herein are those of the authors' and do not necessarily reflect the opinions of the Board of Governors of the Federal Reserve System.

# Overview

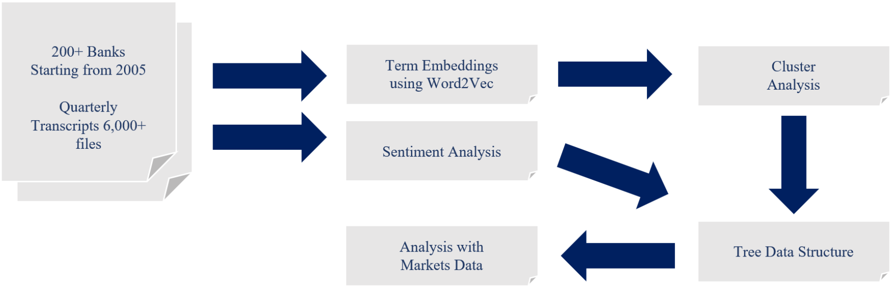
# Motivation

- Large publicly-listed banks have conference calls when they announce quarterly earnings.
  - Which topics are of focus or take up the most time?
  - What is the sentiment associated with various topics?
  - Does this matter for market prices, specifically equity and CDS prices?

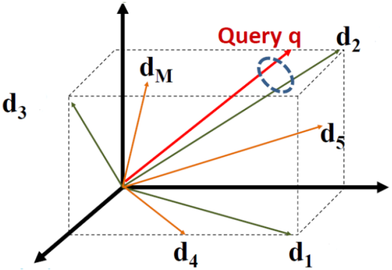
# Literature

- Discussions in earnings conference calls or corporate disclosures are useful in understanding market movements, earnings surprises, or future earnings.
  - Sentiment: Tetlock, Saar-Tsechansky, and Macskassy (2008), Loughran and McDonald (2011), Mayew and Venkatachalam (2012), Price, Doran, Peterson, and Bliss (2012), Jiang, Lee, Martin, and Zhuo (2019), De Aicus, Falconieri, and Tastan (2020)
- **Mostly based on entire body of text/words and dictionary approach**.
- Machine learning approaches.
  - Sentiment: Yang, Siy, and Huang (2020), Lu (2021)
  - Topic: Huang, Lehavy, Zang, and Zheng (2018)
- **We combine tree-structure topic modeling with BERT-based sentiment**.

# Topic Model: Roget Model Overview



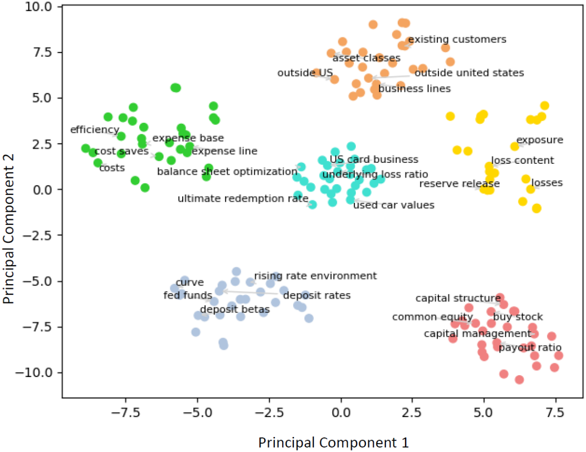
# Topic Model: Word2Vec




Assume you have M expressions (words, paragraphs, slides, documents): Each can be represented by a vector and compared to the other via the vector distance:

If a new expression, or query is presented, it can be compared with the existing expressions and the closest one(s) can be found. **Clusters can be found within the existing vectors.**

# Topic Model: Clusters

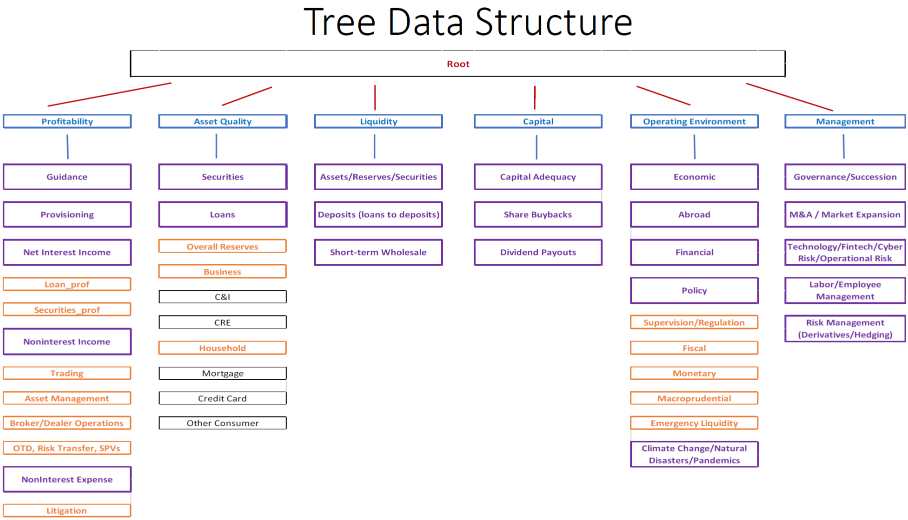



Resulting in interpretable high-level clusters and measurable distance between term representations.

# Topic Model: Tree Data Structure

- Topics arranged in a tree structure starting with the "root"
  - Each node contains its name, Word2Vec representation, topic weights, and average topic sentiment
- Parent nodes have multiple children with a sub-topic relationship
- Topic weight/sentiment averaging is bottom-up: starting with the leaf nodes, the topic sentiment and weights ascend the tree to the root
- The model is therefore a semi-supervised (or weakly supervised) approach

# Topic Model: Tree Data Structure



# Topic Model: Tree Data Structure

**Dominant Topics *Down* the Tree – SVB Financial Group (SIVB Q4 2022 Earnings Conference Call Transcript January 19, 2023, 6:00 PM ET**

| | Topic 1 | Topic 2 |
|---|---|---|
| Yes, Casey. I'll just go to the results of the fourth quarter as an indicator of why that statement makes sense. We're looking at venture deployment in the quarter, 35 billion or so think of that is kind of an annualized run with 120 billion to 140 billion venture deployment in the quarter, from a balance sheet perspective on balance sheet. While we did see the decline in deposits, it was much lower than what we saw in the third quarter. And the reason for that gets to what Greg mentioned, as well as Mike, where we're seeing that lower level of cash burn. So even on a much slower venture deployment number, call it in the mid $30 billion range, we started to see that on balance sheet deposit when we look at cash burn versus inflows get to a much more normalized level.<br><br>So that I think is an indicator with cash burn continuing to slow based on what Mike and Greg just said, that we can get back without going to the 2021 deployment level, to not just deposit of being at the same level, but the potential for deposit growth. | Liquidity | Deposits |

# Sentiment Analysis using FinBERT

- Yang et al. 2020
- Same architecture as BERT, but not BERT based
- Pre-trained on 2.5 Billion tokens (more than BERT):
  - Corporate reports (25 years)
  - Earnings call transcripts (15 years)
  - Analyst reports (13 years)
- Fine-tuned on Analyst Tone - labeled sentences from financial analyst reports
- Results on several financial sentiment datasets

## *Negative* Sentiment Example – JPMorgan Chase (JPM) Q3 2020 Earnings Call Transcript October 13, 2020, 8:30AM ET

*Question:* I may have missed this, I had to jump off for a minute on the call, but can you give us some color? As you've spoken very well about what's going on in the consumer credit area, but when you go into the commercial side of the business, can you share what the sectors that you're seeing the biggest challenges? And can you give us some color on the rerating process that you're going through on those credits that are in trouble today and what kind of deterioration you're seeing in those specific credits in terms of possibly write-downs or revaluations or if it's collateral, like in a commercial real estate loan?

*Answer:* Sure. So, the sectors, I think, are ones that you would expect airlines, lodging, restaurants, other T&E, real estate, oil and gas. And those continue to be the sectors under the most pressure. When you look at downgrades here in the third quarter -- or not here in the third quarter -- in the third quarter, we saw downgrades slow a bit because in the second quarter, we saw significant downgrades just on the increased level of debt that companies were taking on. So, we saw downgrades slow a bit in the third quarter, but we do expect downgrades to continue, particularly in real estate. And then elsewhere, in wholesale, I would say CEO sentiment is guarded, but constructive.

# Market Data Sources

- **SNL Database**
  - EPS consensus estimates and actual values
- **FRB-FAME Database**
  - CDS and equity prices
- **FRB-Fitch Database**
  - Control variables–ex: total assets, net income, gross loans, etc.

# Data: Summary Statistics

| Variable Name | Obs. (GSIBs) | Average | Std. Dev. | Minimum | Maximum |
|---|---|---|---|---|---|
| Root | 2,537 | 0.145 | 0.147 | -0.530 | 0.756 |
| Profitability | 2,537 | 0.142 | 0.180 | -0.947 | 0.974 |
| Asset Quality | 2,537 | 0.073 | 0.271 | -0.999 | 1 |
| Liquidity | 2,537 | 0.110 | 0.266 | -0.994 | 1 |
| Capital | 2,537 | 0.109 | 0.216 | -0.922 | 1 |
| Operating Environment | 2,537 | 0.146 | 0.189 | -0.457 | 1 |
| Management | 2,537 | 0.222 | 0.244 | -0.994 | 0.999 |
| Abnormal returns: Equity: 60 days | 2,175 | -0.001 | 0.038 | -0.441 | 0.153 |
| Abnormal returns: Equity: 120 days | 2,175 | -0.001 | 0.037 | -0.407 | 0.161 |
| CDS (4-Quarter Difference) | 2,147 | -0.001 | 0.046 | -0.402 | 0.488 |

# Calculating Equity Abnormal Returns

$$r_{i,t} - rf_t = \alpha + \beta_1 Mkt_t + \beta_2 SMB_t + \beta_3 HML_t$$

- For t-121 through t-1 to calculate betas
- Where:
  - *Mkt*, *HML*, and *SMB* are Fama-French daily factors
  - *rf* is the risk-free rate

$abnormal\ return = r_{i,t} - \hat{r}_{i,t}$ on earnings release date t

# Calculating CDS Movement

$$CDS_{i,t} = \alpha + \beta_1 CDS_{s,t}$$

- Where CDSs,t is the change in sovereign CDS for the firm i's country

# Sentiment Regressions

$$y_{i,t} = \Theta_i + \lambda_t + \beta_1 S_{c,i,t} + \epsilon_{i,t}$$

for entity i on earnings release quarter t to calculate betas

- Where:
  - $y_{i,t}$ is the abnormal returns or CDS spread for entity i on earnings release date t
  - $\theta_i$ is the fixed effect for entity i
  - $\lambda_t$ is the time fixed effect at quarter date t
  - $S_{c,i,t}$ is the sentiment score for a given category c, for entity i on earnings release quarter t
    - Seven categories: root, profitability, liquidity, asset quality, capital, management, operating environment
    - Run each category of sentiment separately, then all at once
    - Can add a dummy that indicates whether earnings missed analysts' forecasts

# GSIB Abnormal Equity Returns and Sentiment

**Sentiment | -999 as 0 | GSIB ONLY**

| | DV: Abnormal Return for Equity Prices - 60 Day | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Root | 0.045$^{***}$ | | | | | | | | | |
| | (0.007) | | | | | | | | | |
| Profitability | | 0.029$^{***}$ | | | | | | | | 0.020$^{***}$ |
| | | (0.006) | | | | | | | | (0.007) |
| Liquidity | | | 0.008$^{**}$ | | | | | | | 0.003 |
| | | | (0.004) | | | | | | | (0.004) |
| Capital | | | | 0.006 | | | | | | 0.002 |
| | | | | (0.004) | | | | | | (0.004) |
| Asset Quality | | | | | 0.014$^{***}$ | | | | | 0.009$^{**}$ |
| | | | | | (0.004) | | | | | (0.004) |
| Management | | | | | | 0.011$^{***}$ | | | | 0.002 |
| | | | | | | (0.004) | | | | (0.005) |
| Operating Environment | | | | | | | 0.016$^{***}$ | | | 0.002 |
| | | | | | | | (0.006) | | | (0.006) |
| ROA | -0.001 | -0.001 | -0.0004 | -0.0004 | -0.0005 | -0.0004 | -0.0005 | 0.00004 | | -0.001 |
| | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | | (0.0004) |
| MISSED EST | -0.016$^{***}$ | -0.016$^{***}$ | -0.018$^{***}$ | -0.018$^{***}$ | -0.018$^{***}$ | -0.017$^{***}$ | -0.018$^{***}$ | | -0.017$^{***}$ | -0.016$^{***}$ |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | | (0.002) | (0.002) |
| Observations | 2,175 | 2,175 | 2,175 | 2,175 | 2,175 | 2,175 | 2,175 | 2,490 | 2,215 | 2,175 |
| $R^2$ | 0.073 | 0.063 | 0.050 | 0.048 | 0.057 | 0.051 | 0.052 | 0.00001 | 0.042 | 0.068 |
| Adjusted $R^2$ | -0.482 | -0.498 | -0.520 | -0.524 | -0.509 | -0.517 | -0.516 | -0.517 | -0.522 | -0.496 |
| F Statistic | 35.888$^{***}$ (df = 3; 1359) | 30.677$^{***}$ (df = 3; 1359) | 23.816$^{***}$ (df = 3; 1359) | 22.611$^{***}$ (df = 3; 1359) | 27.154$^{***}$ (df = 3; 1359) | 24.544$^{***}$ (df = 3; 1359) | 24.992$^{***}$ (df = 3; 1359) | 0.009 (df = 1; 1641) | 61.584$^{***}$ (df = 1; 1393) | 12.441$^{***}$ (df = 8; 1354) |
| Significance Levels | | | | | | | | | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

# Abnormal Equity Returns and Sentiment

**Sentiment | -999 as 0 | GSIB ONLY | Profitability: Children**

| | DV: Abnormal Return for Equity Prices - 60 Day |
|---|---|
| General.Profitability.and.Guidance | -0.001 |
| | (0.003) |
| Provisioning | 0.007$^{*}$ |
| | (0.004) |
| Net.Interest.Income | 0.001 |
| | (0.004) |
| Noninterest.Income | 0.018$^{***}$ |
| | (0.004) |
| Noninterest.Expense | -0.0001 |
| | (0.004) |
| roa_diff_ann | -0.001 |
| | (0.0004) |
| miss_est | -0.016$^{***}$ |
| | (0.002) |
| Observations | 2,175 |
| $R^2$ | 0.063 |
| Adjusted $R^2$ | -0.503 |
| F Statistic | 13.103$^{***}$ (df = 7; 1355) |
| Significance Levels | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

# CDS Spread Movement and Sentiment

**Sentiment | -999 as 0 | GSIB ONLY**

| | DV: Abnormal Return for CDS Prices (% Change) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Root | -0.017** | | | | | | | | |
| | (0.008) | | | | | | | | |
| Profitability | | -0.007 | | | | | | | -0.006 |
| | | (0.006) | | | | | | | (0.007) |
| Liquidity | | | 0.004 | | | | | | 0.007* |
| | | | (0.004) | | | | | | (0.004) |
| Capital | | | | -0.016*** | | | | | -0.015*** |
| | | | | (0.005) | | | | | (0.005) |
| Asset Quality | | | | | 0.003 | | | | 0.006 |
| | | | | | (0.004) | | | | (0.004) |
| Management | | | | | | -0.006 | | | -0.003 |
| | | | | | | (0.004) | | | (0.005) |
| Operating Environment | | | | | | | -0.010* | | -0.007 |
| | | | | | | | (0.006) | | (0.007) |
| Tier1 Capital Ratio | -0.002* | -0.002* | -0.002* | -0.002* | -0.002* | -0.002* | -0.002* | -0.002* | -0.002* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Observations | 2,147 | 2,147 | 2,147 | 2,147 | 2,147 | 2,147 | 2,147 | 2,147 | 2,147 |
| $R^2$ | 0.006 | 0.004 | 0.003 | 0.010 | 0.003 | 0.004 | 0.005 | 0.002 | 0.015 |
| Adjusted $R^2$ | -0.602 | -0.605 | -0.606 | -0.595 | -0.607 | -0.605 | -0.603 | -0.606 | -0.592 |
| F Statistic | 3.943** (df = 2; 1332) | 2.385* (df = 2; 1332) | 2.259 (df = 2; 1332) | 6.883*** (df = 2; 1332) | 1.904 (df = 2; 1332) | 2.701* (df = 2; 1332) | 3.179** (df = 2; 1332) | 3.293* (df = 1; 1333) | 2.944*** (df = 7; 1327) |
| Significance Levels | *p<0.1; **p<0.05; ***p<0.01 | | | | | | | | |

# Concluding Remarks

- Bank equity prices are sensitive to information conveyed in the profitability part of the bank earnings calls
- CDS premiums are sensitive to what is revealed in the capital discussions
- Application: ROGET

# How does the model handle semantic shift and uncommon words?

Examples: COVID-19, Climate, Environment, Weathering (the economic or environment storm?)

- Add to the mapping of words and phrases (semi-supervised)
- For each new block of text, the embedding model is updated to include the current data. The tree is re-measured in the bottom-up fashion
- A weighting function is used in every re-measurement of the tree to mitigate against the distance becoming skewed due to larger/smaller amounts of text in a given block

Thank you!

Contact Information


- Dylan Saez: dylan.c.saez@frb.gov
- Seung Lee: seung.lee@frb.gov