Constructing High-frequency and Thematic Economic Sentiment Indicators from News Articles

ALAN CHESTER ARCIN, CARMELITA ESCLANDA, CHELSEA ANNE ONG, ROSSVERN REYES

3RD IFC AND BANK OF ITALY WORKSHOP ON "DATA SCIENCE IN CENTRAL BANKING: ENHANCING THE ACCESS TO AND SHARING OF DATA", OCTOBER 17-19, 2023

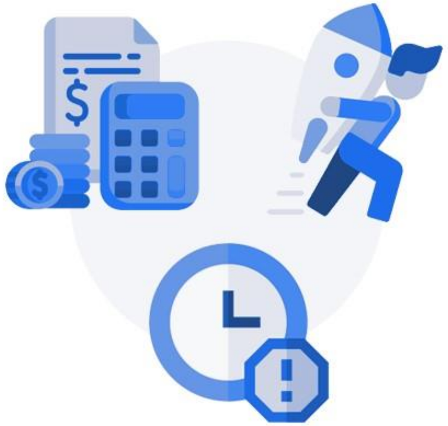
* The views expressed herein are those of the authors' only and do not necessarily reflect those of the Bangko Sentral ng Pilipinas

Outline

1. Motivation
2. Overview, Data Collection and Pre-processing
3. Methodology
4. Model Evaluation
5. Results – Sentiment Indices
6. Topic Modelling
7. Results – Topic modeling
8. Key Takeaways and Future work

Motivation

Sentiment influences economic agents' decisions



Survey-based measures of sentiment are **costly, tedious, and published with a delay**

News-based indices can be **produced at a higher frequency**, allowing for more up to date assessment at a lower cost

Overview

Web scraping

Data cleaning

Sentiment analysis

NSI

Data Collection and Pre-processing

Data were collected from multiple local news sources. The sections of interest were limited to economy, banking, finance and related fields. Text from the articles are pre-processed and further analyzed.

Pre-processing steps include:

- Removal of extra spaces in the text
- Case normalization
- Removal of punctuations and numbers
- Unicode conversion
- Tokenization
- Vectorization

Methodology

Dictionary based method involves a pre-defined list of keywords from some of the most used dictionaries in finance and well-known general dictionaries:

- General Inquirer (GI Lexicon)
- Loughran-McDonald Sentiment Dictionary (LM Lexicon)
- Financial Stability Sentiment Dictionary (FD Lexicon)
- Hu and Liu Opinion Lexicon (HL Lexicon)
- Valence Aware Dictionary for Sentiment Reasoning (VADER)

Using genetic algorithm, we retrieved combinations of keywords that maximizes the number of classified sentences from ~3000 manually labelled sentences (PH Lexicon).

Methodology

Machine learning based method utilizes two pretrained models to predict sentiment of text:

- BERT
  - ML model specifically trained using Wikipedia and BooksCorpus for common language tasks

- FinBERT
  - Built by further training BERT using financial text such as 10-K and 10-Q Corporate Reports, analyst reports and fine tuning it for financial sentiment classification

Model Evaluation

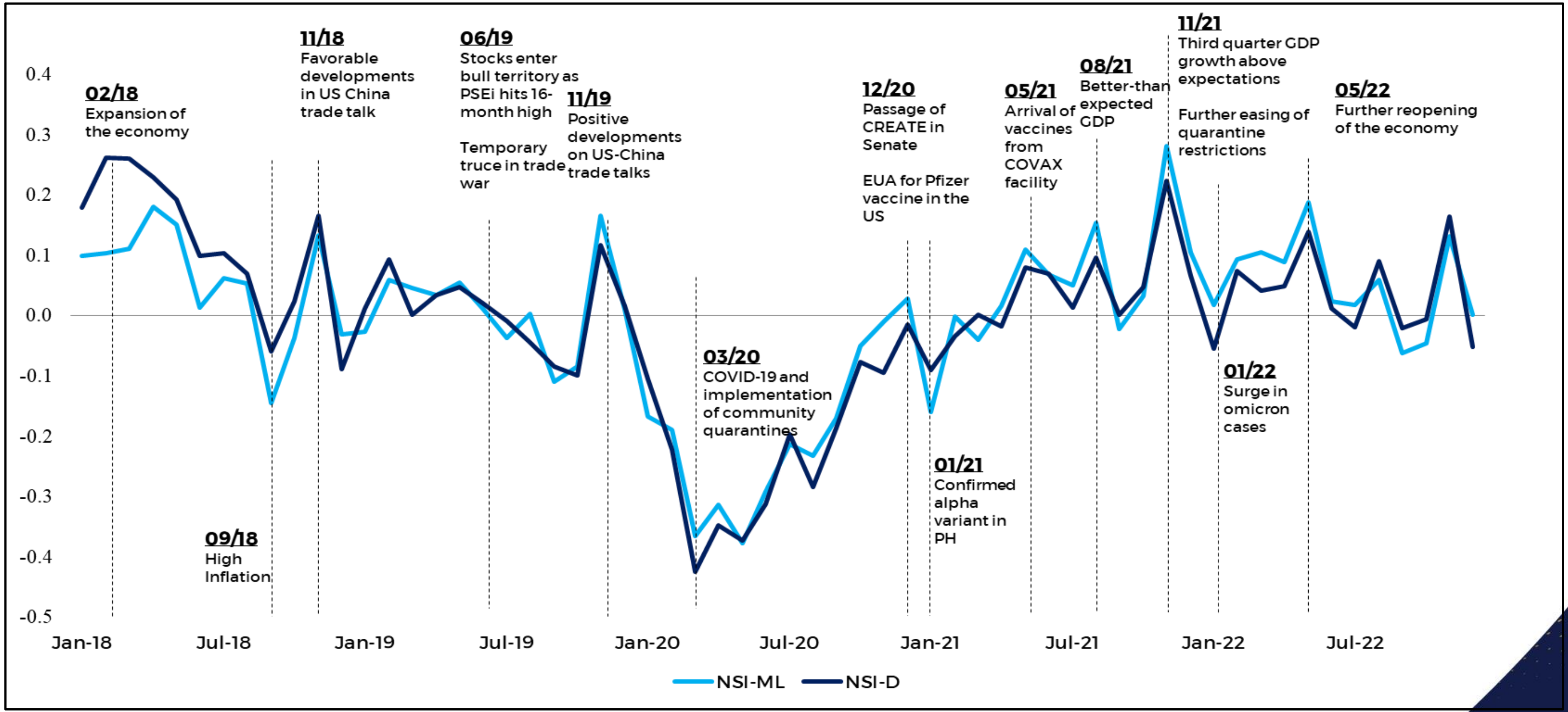
Both dictionaries and machine learning models were evaluated against ~3000 manually labelled sentences sampled from online news articles

Lexicons	Accuracy	Macro F1
PH Lexicon	0.66	0.64
LM Lexicon	0.52	0.46
FD Lexicon	0.51	0.51
VADER	0.51	0.42
HL Lexicon	0.50	0.46
GI Lexicon	0.45	0.42
Machine Learning Models	**Accuracy**	**Macro F1**
FinBERT	0.66	0.63
BERT	0.58	0.45

Accuracy measures the fraction of the model's correct predictions over total predictions. Meanwhile, macro F1 score combines both precision (a measure of the proportion of true positives that were correct) and recall (a measure of the proportion of true positives that were actually predicted).

Results

Estimates of NSIs strongly co-move with key economic events.



Classification: GENERAL

Results

Correlation analysis suggest co-movements between NSIs and selected economic indicators

Model	PSEi	PMI	BES	CES
NSI-D	0.61	0.62	0.49	0.38
NSI-ML	0.53	0.58	0.43	0.28

The Business Expectations Survey (BES) is one of the quarterly surveys conducted by the Bangko Sentral ng Pilipinas (BSP) which gathers information about entrepreneurs' views on the general business situation in their own industry and on the national economy. Likewise, the Consumer Expectations Survey (CES) captures the economic outlook of consumers as an indication of the country's future economic condition.

Topic Modeling and Correlation Analysis

Topic Modeling involves unsupervised machine learning techniques to **cluster text based on similarities**.

We considered the two of the commonly used techniques in topic modelling:

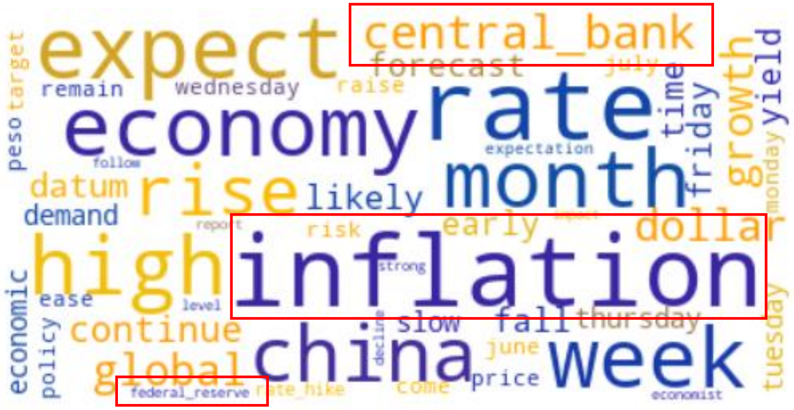
- **Non-Negative Matrix Factorization (NMF)**
- Latent Dirichlet Allocation (LDA)

Using the coherence score, we determined **12** to be the optimal number of topics. We then group the articles according to topics, compute the NSI for each and correlate these with selected economic indicators.

Results

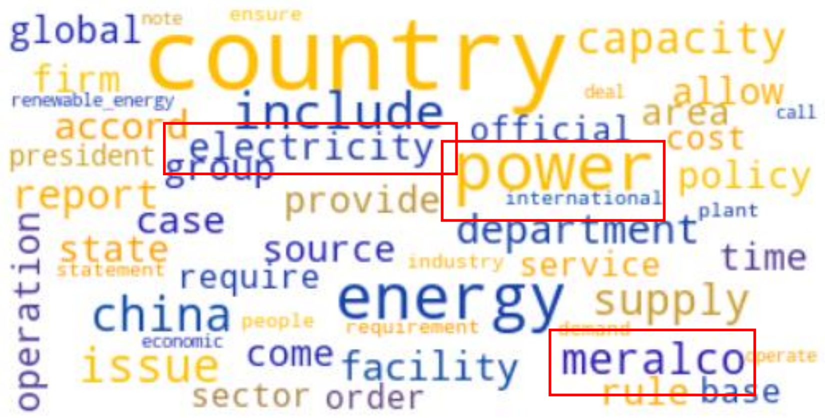
Word clouds of selected topics

Banking



Monetary Policy

Company Earnings



Energy

Agriculture

Results

Correlation analysis show comovements of selected topics against monthly indicators PSEi and PMI…

Topic	PSEi	PMI
Company Earnings	0.72	0.53
Services	0.45	0.47
Trade	0.44	0.63
Banking	0.41	0.49
Government Projects	0.44	0.35

As well as against quarterly indicators BES, CES and GDP growth rate

Topic	BES CI	CES CI	GDP
Company Earnings	0.74	0.79	0.81
Infrastructure	0.60	0.62	0.70
Trade	0.48	0.36	0.78

Key Findings and Future Works

Key Takeaways:

- NSIs are found to **match expected sentiment for key economic events**
- NSIs showed potential in **complementing existing surveys** (BES and CES) and providing **timely measurement of market sentiment**
- Topic modeling can **help uncover specific news themes** that may have more predictive content for specific key economic variables

Future Works:

- Enhance NSI-D (e.g., Add more linguistic rules, Part-of-Speech Tagging)
- Retrain FinBERT using the manually labelled dataset for NSI-ML
- Complement with other ML models (e.g., LM Lexicon + SVM, FinBERT + Random Forests, etc.)

Sources

- Stone, P. J., Dunphy, D. C., & Smith, M. S. (1966). The general inquirer: A computer approach to content analysis. M.I.T. Press.
- Loughran, Tim and McDonald, Bill, When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks (March 4, 2010). Journal of Finance, Forthcoming, Available at SSRN: https://ssrn.com/abstract=1331573
- Correa, Ricardo, Keshav Garud, Juan M. Londono, and Nathan Mislang (2017). Sentiment in Central Banks' Financial Stability Reports. International Finance Discussion Papers 1203.
- Zhang, L., Liu, B. (2017). Sentiment Analysis and Opinion Mining. In: Sammut, C., Webb, G.I. (eds) Encyclopedia of Machine Learning and Data Mining. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7687-1_907
- Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805. Retrieved from https://arxiv.org/abs/1810.04805
- Huang, Allen and Wang, Hui and Yang, Yi, FinBERT - A Large Language Model for Extracting Information from Financial Text (July 28, 2020). Contemporary Accounting Research, Forthcoming, Available at SSRN: https://ssrn.com/abstract=3910214 or http://dx.doi.org/10.2139/ssrn.3910214

Thank you!

* The views expressed herein are those of the authors' only and do not necessarily reflect those of the Bangko Sentral ng Pilipinas