

How to Interpret an Economic Index? Generating Reports with Topic Sentiment Analysis

Kazuhiro Seki

Konan University

seki@konan-u.ac.jp



Kazuhiro Seki, Ph.D.

- Professional Experience

- Professor in the Faculty of Intelligence and Informatics, Konan University
- Senior Research Fellow at Asia Pacific Institute of Research
- Research Fellow at Research Institute for Economics and Business Administration, Kobe University

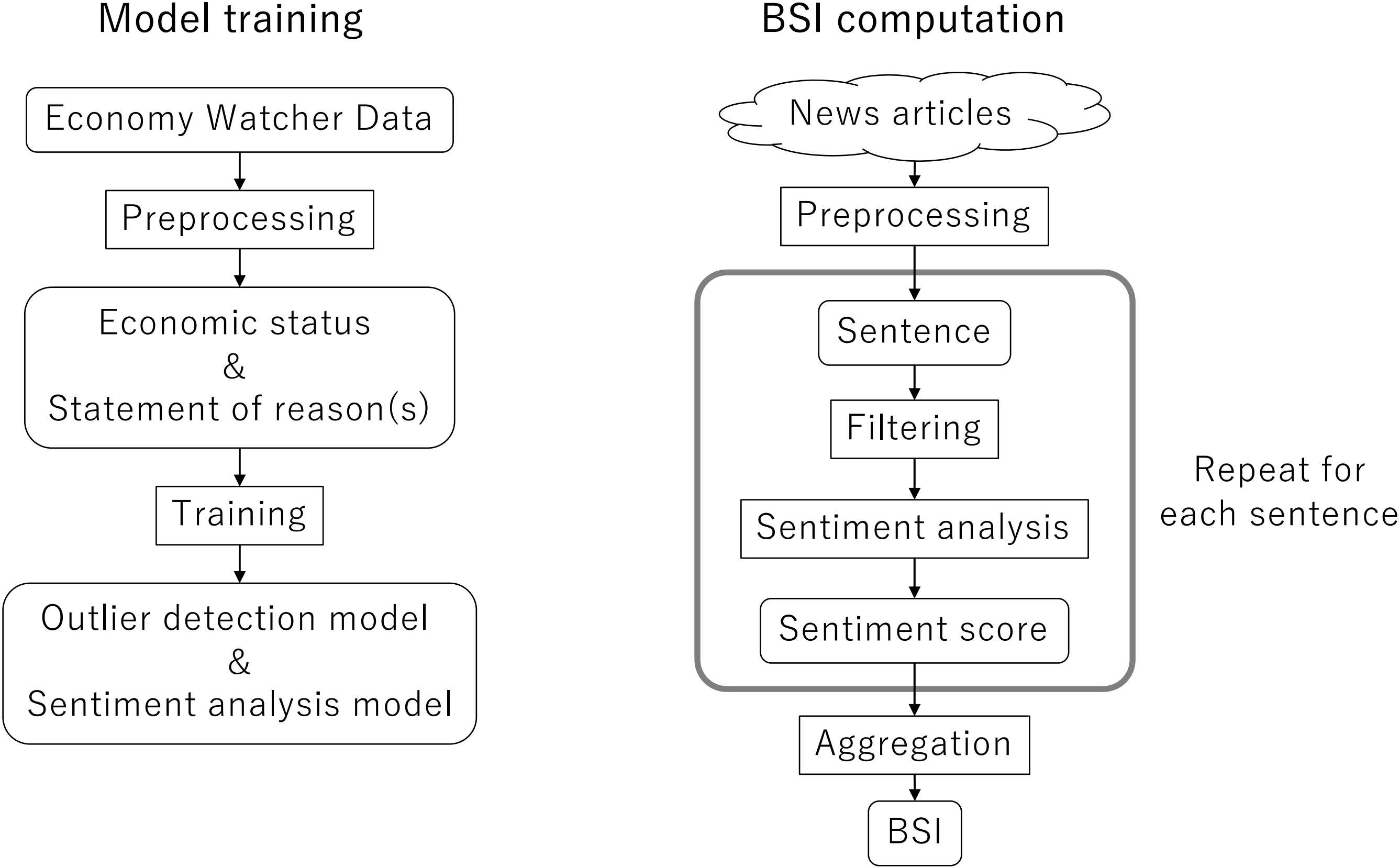
- Research Areas and Activities

- Text analytics, data mining, machine learning, and their applications to real-world problems
- Coauthored more than 70 publications in leading conferences and journals, including SIGIR, WSDM, CIKM, IJCAI, IP&M, JASIST, and Machine Learning

Background

- Using news articles, we've been working on nowcasting an economic index, Business Sentiment Index (BSI)
 - BSI measures economic confidence
 - Traditional BSIs are based on surveys, which suffer from high costs and delays

News-based BSI Nowcasting System




Adapted from: Seki et al.. News-based Business Sentiment and its Properties as an Economic Index. *IP&M*, Vol. 59, No. 2, 102795, 2022.

News-based BSI Nowcasting System



Motivation and Goal

- News-based BSI improves timeliness and cost effectiveness, which however lacks interpretability
 - Why is BSI -0.17 this month?
- 
- Goal: To generate economic reports rationalizing the values of BSI
 - Topic-based two-stage summarization

Methods: Overview

- News articles used for estimating BSI are reused for generating economic reports
- Two-stage summarization:
 1. Topic-based summarization
 2. Overall summarization
- Use a topic model to identify latent topics and an LLM to generate par topic and overall summaries

Methods: Summarization Process

1. Data Preparation:

- Collect news articles, estimate sentiment scores by a finetuned BERT (already done)

2. Topic Identification:

- Apply a neural topic model to extract k latent topics

3. Topic-Based Summarization:

- Extract n key sentences per topic based on topic sentiment scores $s_i \cdot p(t|d_i)$
- Generate topic label and topic summary using an LLM

4. Overall Summarization:

- Merge topic summaries for a comprehensive report

Experiments: Settings

- Data: Sankei news articles (October 2024)
 - News-based BSI = -0.17
- Topic model: FASTopic [Wu+, NeurIPS 2024] with paraphrase-multilingual-mpnet-base-v2 for embedding
- LLM: GPT-4o mini for summarization
- Parameters:
 - $k = 5$ (number of topics)
 - $n = 100$ (sentences per topic)

First Prompt to Generate a Topic Summary

The news articles listed below belong to the same topic as determined by topic analysis. The number at the end of each sentence indicates its impact on the economy. For each topic, generate an appropriate and concise label, and summarize the key points, particularly from the perspective of economic impact, in about three bullet points. Output the results in JSON format according to the structure below.

Output Format

```
{
  "label": "Topic label",
  "points": [
    "Sentence indicating a key point",
    "Sentence indicating a key point",
    "Sentence indicating a key point"
  ]
}
```

News Sentences

- 1st news sentence (topic-sentiment score)
- 2nd news sentence (topic-sentiment score)
- ...

Experimental Results

- Five key topics identified by FASTopic (n key sentences are omitted)
 - Crime and Security Issues (-0.64)
 - Economic Performance Overview (-0.19)
 - Despair and Isolation (-0.41)
 - Energy and Infrastructure Development (-0.20)
 - Security and Geopolitical Tension (-0.58)

topic-sentiment
distribution
 $E_{d_i}[s_i \cdot p(t|d_i)]$

Second Proprompt to Generate a Topic Summary

Below is a summary of recent events for each of the k topics that are considered to have a significant impact on the economy. Each topic has been assigned an appropriate label. Generate an overall summary of approximately 300 characters. Output the results in JSON format according to the structure below.

Output Format

```
{  
  "summary": "Summary of the overall content"  
}
```

Reports by Topic (JSON Format)

[k outputs from the previous step]

Our Topic-based Summary

Recent events have revealed a troubling intersection of **crime**, **economic performance**, and **geopolitical tensions** affecting public sentiment. **Rising violent crime and failures in police oversight have heightened societal fears**, while **economic data shows strong sales juxtaposed with declining consumer confidence due to rising costs**. Alongside this, feelings of despair and isolation are on the rise, impacting mental health and productivity. In the face of mounting energy demands driven by digital advancements, there's a pressing need for sustainable infrastructure, notably nuclear power, despite the historical challenges faced. Finally, Japan's complex security environment calls for robust foreign policies as geopolitical tensions escalate, particularly due to the implications of Russia's actions and China's assertiveness, threatening both regional and global stability. [color-coded by topics]

Baseline: One-step Summary from 500 High Sentiment Sentences

Summarize the excerpts from newspaper articles listed below from the perspective of their impact on the economy. The number at the end of each article indicates the degree of its impact on the economy. Each sentence is part of an independent article, so there is generally no connection between sentences.

Output the result in JSON format following the structure below.

Output Format

```
{  
  "summary": "A summarized text covering all excerpts"  
}
```

News Sentences

- 1st news sentence (topic-sentiment score)
- 2nd news sentence (topic-sentiment score)
- ...

Baseline Summary

The news articles present a mixed picture regarding economic impacts. On one hand, some reports indicate a **significant increase in consumer activity and investment, such as in advertising and sales of certain products**, along with **a strong recovery in tourism from abroad**. Conversely, there are alarming signs of **economic decline, with a dramatic rise in bankruptcies, particularly among small businesses, and growing concerns over rising costs** and declining consumer confidence. Overall, while some sectors appear to thrive, many are struggling under pressures of financial instability and societal challenges. [color-coded by topics]

Comparison with Baseline

- Our Topic-based Approach:
 - More comprehensive summaries covering wider topics
 - Captures geopolitical, societal, and economic factors

Conclusions and Future Work

- Proposed a topic-centered summarization approach
- Generated an economic report which helps us understand the context of the news-based BSI
- Demonstrated improved coverage compared to a baseline
- Future Work:
 - Integrate the approach into our web-based business sentiment analysis system (<https://sapis.apir.or.jp>)