COSE474-2024F: Final Project Proposal
"Predicting Goods Price Changes Based on Sentiment Analysis of News Headlines"

Mirza Syahmi Bin Afandi

1. Introduction

Economic forecasting is critical for consumers, businesses, and policymakers, especially in the context of price fluctuations for essential goods. Traditional models rely heavily on historical data and market variables, often overlooking the potential influence of current events as captured by media. This project aims to explore how real-time sentiment analysis of news headlines can predict changes in goods prices.. By incorporating public sentiment as expressed in media, this project offers a novel approach to forecasting price dynamics in the marketplace.

2. Problem definition & challenges

The core problem is to predict monthly changes in consumer goods prices using sentiment scores derived from news headlines. This project will develop a deep learning-based model that employs sentiment analysis to understand and forecast price fluctuations, reflecting the influence of economic conditions and global events reported in the media.
Challenges include accurately capturing sentiment from diverse and unstructured news data, building a meaningful link between extracted sentiment and actual price changes, and aligning news headlines with price data on a time scale that captures relevant market reactions without introducing noise from unrelated events.

3. Related Works

(Nemes & Kiss, 2021) introduced the application of sentiment analysis on stock news headlines to predict stock value changes.
(Putra et al., 2023). proposed a machine-learning-based approach to predict tourist arrival trends using online news headlines and previous tourist arrival data.

4. Datasets

I will use publicly available monthly consumer goods prices from the Department of Statistics Malaysia. For the news headline, I will be using The Star (a well-known Malaysian news website and newspaper) News Article dataset from Kaggle.

5. State-of-the-art methods and baselines

State-of-the-art methods for predicting goods price changes primarily use time-series models like ARIMA and LSTM. Pre-trained sentiment analysis models such as BERT effectively interpret text but aren't applied to price forecasting. I will use these models as baselines to assess the impact of integrating news sentiment into goods price prediction.

6. Schedule

Week 1: Data collection and preprocessing. Week 2: Implement sentiment analysis using pre-trained models BERT on the news headlines. Week 3: Develop baseline models (ARIMA, LSTM) for goods price prediction based on historical data. Week 4: Integrate sentiment analysis results into the predictive model. Week 5: Model training, testing, and performance evaluation; tune hyperparameters for optimal results. Week 6: Final validation, error analysis, and finalize project report.

References

Nemes, L. and Kiss, A. Prediction of stock values changes using sentiment analysis of stock news headlines. *Journal of Information and Telecommunication*, 5(3):375–394, 2021.

Putra, I. P. E. S., Lestari, D., and Tunjungsari, K. R. A machine-learning-based approach for tourist-arrival trend prediction. *Journal of Business on Hospitality and Tourism*, 9(2):212–233, 2023.