

# Twitter Sentiment Analysis on Fuel Inflation Issue in Indonesian using Random Forest, Naïve Bayes, and Support Vector Machine

Muhamad Rikbal Ikhsani<sup>1</sup>, Bambang Ari Wahyudi<sup>2</sup>, Irma Palupi<sup>3</sup>

<sup>1,2,3</sup>Fakultas Informatika, Universitas Telkom, Bandung

<sup>4</sup>Divisi Digital Service PT Telekomunikasi Indonesia

[muhamadrikbal@students.telkomuniversity.ac.id](mailto:muhamadrikbal@students.telkomuniversity.ac.id), [bambangari@telkomuniversity.ac.id](mailto:bambangari@telkomuniversity.ac.id),

[pembimbing2@telkomuniversity.ac.id](mailto:pembimbing2@telkomuniversity.ac.id)

## 1. Introduction

### 1.1 Background of The Study

As an archipelagic country with a large population, Indonesia has high transportation needs. The transportation sector is critical to supporting people's basic needs, such as clothing, food, and shelter. The transportation sector is very closely related to energy needs, with 90% of energy in fuel oil (BBM) [1]. Currently, the public often responds and criticizes political and public leaders through social media such as Twitter. Twitter is one of the social media that has a retweet feature that users can use to re-upload information or tweets, which allows the dissemination of information on Twitter social media to be faster [2]. When people respond to fuel price increases on Twitter, there are always many pros and cons.

The exponential increase in the amount of information available on social media has made sentiment analysis increasingly important [3]. Sentiment analysis is a technique used to analyze text and determine the subjectivity of opinions expressed in the text, such as in reviews or tweets [4]. This makes it possible to tell whether the general opinion on a particular topic is positive, negative, or neutral. In government performance, sentiment analysis can determine public opinion about the current government performance. If the general opinion is negative, it may indicate that the government needs to improve their performance in certain areas. If public opinion is positive, it can indicate that the government is performing well. Sentiment analysis can be used as an aid in making decisions and determining policies by policymakers.

Many classification methods can be used in sentiment analysis, including K-Nearest Neighbor, Naïve Bayes, Support Vector Machine, Random Forest, etc. One popular method used is the Support Vector Machine (SVM) classification. In L. Mandloi and R. Patel's research, researchers obtained good accuracy in the SVM method for classifying sentiments in the text [7]. SVM is also used in the classification of Twitter comments by MA Al-Ghonaim entitled "Deep Learning and SVM-Based Sentiment Analysis of Twitter Data", with high accuracy results [8]. Other popular methods are Naïve Bayes and Random Forest. In the research of Indra Budi and Dian Arianto [5], entitled Aspect-based Sentiment Analysis on Indonesia's Tourism Destinations Based on Google Maps User Code-Mixed Reviews (Study Case: Borobudur and Prambanan Temples), the Random Forest and Naïve Bayes methods were used for analysis sentiment on tourist destination reviews on Google Maps. Then, in research conducted by Negis et al. [6] entitled "Sentiment Analysis Using A Random Forest Classifier On Turkish Web Comments", the Random Forest method is also used for sentiment analysis on comments on Turkish-language websites.

The purpose of this research is to analyze sentiment on Twitter comments regarding the issue of rising fuel prices using the three methods that have been discussed. This research will present a comparison of the effectiveness of the three methods in sentiment classification. By comparing different methods, the writer will get results regarding the ability of different approaches to distinguish sentiment on Twitter comments about the fuel price hike issue.

### 1.2 Topic and Limitation

This final project focuses on sentiment analysis in the case of rising fuel prices using the Support Vector Machine, Naïve Bayes and Random Forest methods. This study has limitations, namely: (1) The dataset used is a collection of 2961 tweets in Indonesian from 2017-2022; (2) The data is divided into two categories of sentiment, namely positive and negative; (3) The selected discussion is only limited to the increase in fuel prices.

### 1.3 Purpose

This study aims to create naive Bayes algorithms, support vector machines and random forests to conduct sentiment analysis in cases of rising fuel prices to obtain the best algorithm.

### 1.4 Writing Organization

The writing organization in this study is structured as follows. The first section consists of an introduction. The second section contains an explanation of related studies from this research. The third section contains an explanation of the method used and the application of the system that has been built. The fourth section contains the test results and analysis. Then the fifth section contains the conclusions of this study.