

1. Introduction

Sentiment is an opinion of a person towards a certain entity [1]. Sentiments can be analyzed to determine people's opinions, emotions, views, and attitudes towards something [2]. Sentiments can be found in comments, websites, social media, and blogs, useful as indicators for various purposes [3]. Sentiment is beneficial for companies, businesses, and governments, but taking advantage of it requires text mining and sentiment analysis [4].

Based on the Theory Efficient Market Hypothesis (EMH) states that all information in the market reflects the price of the stock [5]–[7]. Information that is widely available and can be accessed by anyone makes stock price movements always occur efficiently and can be analyzed. A study conducted by Sul et al. (2014) found a relationship between tweet sentiment from a company and the movement of the company's shares [8]. One of the cases that have occurred is the case of the lost balance of Bank Mandiri customers in 2019. This case affected 10% of Bank Mandiri customers with losses estimated at 10 billion rupiahs and caused negative sentiment towards Bank Mandiri [9]. Therefore, sentiment analysis on issues and news on social media is carried out because it can affect a stock's price movements.

This study aims to determine the correlation of sentiment from social media and news media to the movement of Bank Mandiri shares. An NBSVM classification model was built with the Word2Vec feature expansion. Then, the accuracy was compared before and after applying Word2Vec to the model and the features that affect the model's accuracy. This classification method is used because it can work well on small data such as document pieces, as well as a full-length documents [10]. This study uses Term Frequency - Inverse Document Frequency (TF-IDF) to convert the data into a vector. Word2Vec has the advantage of seeing the semantic relationship between words [11]. Word2Vec is used to reduce vocabulary errors in sentiment analysis using word embedding [12]. The result of this study is expected to create a model that can analyze sentiment regarding the movement of Bank Mandiri shares to assist investors in making decisions.

The writing organization in this study is as follows. Section 2 will discuss the literature review. Section 3 will discuss system design. Sections 4 and 5 will discuss the result of this research and its conclusions.