

1. Pendahuluan

Pagi pada tanggal 20 Juli 2019, media sosial diramaikan dengan perbincangan mengenai sejumlah nasabah menyampaikan keluhan kesahnya perihal eror-nya layanan Bank Mandiri. Dapat diketahui dari beberapa keluhan yang dilontarkan tersebut, ternyata sebagian nasabah mengeluhkan tentang berkurangnya saldo yang dimiliki sebelumnya berkurang sangat drastis. Dan dilain hal itu, sebagian nasabah juga mengabarkan bahwa saldo yang berada pada tabungannya tersebut bertambah secara tiba-tiba [1]. Melalui sosial media twitter para nasabah menyampaikan keluhan kesahnya perihal eror-nya layanan Bank Mandiri tersebut. Dapat dikutip pada tweet dari salah satu pemilik akun twitter dengan nama user @agung_sampurna1 “ Bank Mandiri bagaimana ini saldo saya habis 600 ribu dan jadi berkurang tentunya padahal tidak ada transaksi transfer atau penarikan. “. Dari tweet tersebut dapat diketahui bahwa opini yang dilontarkannya oleh akun twitter dengan nama user @agung_sampurna1 menandakan bahwa layanan sistem Bank Mandiri sedang terjadi eror [2]. Setelah kejadian tersebut, saham Bank Mandiri pada pembukaan perdagangan senin, 22 Juli 2019 turun sebanyak 0,95 persen menjadi Rp. 7.800 untuk perlembar saham. Ketika pada pukul 12.30 WIB penurunan tetap berlanjut hingga mencapai 1,27 persen menjadi Rp. 7.775 untuk perlembar saham [3].

Berdasarkan permasalahan tersebut peneliti akan mencari keterhubungan antara analisis sentimen dengan harga saham Bank Mandiri tersebut. Peneliti akan mengumpulkan data *tweet* dari *Twitter* dan berbagai media berita online perihal eror-nya layanan Bank Mandiri. Dengan data yang dikumpulkan tersebut peneliti akan melakukan tahapan analisis sentimen dengan maksud tujuan untuk menghasilkan suatu sentimen yang berguna dalam mengetahui tingkatan kepuasan nasabah terhadap suatu layanan Bank Mandiri [4].

Pada *Twitter* penulisan *tweet* dibatasi hingga 280 karakter, sehingga pesan yang ditangkap secara singkat dan bahkan dapat menggunakan banyak variasi kata. Dengan batasan penulisan *tweet* yang diberikan oleh *Twitter* maka menyebabkan pengguna menuliskan banyak singkatan kata, istilah pendek, atau kata dengan salah eja, dan penggunaan kata-kata gaul pada internet. Penggunaan variasi kata tersebut memungkinkan ketidakcocokan kosakata dan sulit untuk memahami terhadap *tweet* tersebut [5]. Tujuan dari penelitian ini adalah untuk memfokuskan bagaimana cara mereduksi kosakata yang kurang cocok dengan menggunakan *word embeddings*. Peneliti menggunakan ekspansi fitur dengan model *FastText*. Dengan tujuan untuk mempelajari kata tengah dari setiap konteks data tekstual yang tersisa atau dari keseluruhan konteks data tekstual, dengan maksud tujuan dapat membantu dalam menangkap arti kata-kata yang lebih pendek tersebut [6]. Pada tahapan klasifikasi peneliti menggunakan klasifikasi *Logistic Regression* untuk menghubungkan antara kedua variabel X dan Y. Dan variabel X yang diprediksi merupakan fungsi dari suatu probabilitas yang menandakan bahwa suatu objek akan berada dalam suatu kategori Y [7].

Rumusan masalah yang dapat diambil berupa implementasi pengaruh sentimen pada saham Bank Mandiri (BMRI) menggunakan ekspansi fitur dengan *FastText* dengan data *corpus pre-trained* berbahasa Indonesia dan klasifikasi *Logistic Regression*, lalu mengukur performansi terbaik dari klasifikasi *Logistic Regression* menggunakan ekspansi fitur dengan menerapkan model *pre-trained* dari *FastText*. Dan menentukan korelasi antara sentimen yang berpengaruh terhadap harga saham BMRI. Data yang digunakan pada penelitian kali ini berasal dari proses *crawling* data berbahasa Indonesia dari laman social media *Twitter* dan media berita online seperti *CNN*, *Detik*, *Liputan 6*, *Stockbit*, *CNBC*, *Tempo*, dan *Youtube*.

Penelitian ini bertujuan untuk mengimplementasikan optimalisasi model yang diterapkan dengan menguji model klasifikasi *Logistic Regression* dengan skenario *split* data, penggunaan jumlah *n-gram*, dengan batasan maksimal *feature*. Dan tahapan skenario yang akan diterapkan berikutnya adalah dengan menguji model *Logistic Regression* dengan implementasi *hyperparameter tuning*, dan skenario yang terakhir dari tahapan optimalisasi model adalah dengan menguji model klasifikasi *Logistic Regression* dengan penerapan ekspansi fitur menggunakan model *FastText*, untuk mengurangi adanya ketidakcocokan kosakata pada suatu data tekstual. Setelah optimalisasi model dilakukan, akan dilakukannya pengujian korelasi menggunakan *Pearson Correlation Coefficient* antara *sentimen* data terhadap harga saham BMRI dengan model terbaik berdasarkan penerapan optimalisasi model. Dengan tujuan untuk mencari pengaruh *sentimen* pada harga saham Bank Mandiri (BMRI).

Pada sub-bagian kedua (Studi Terkait) akan membahas penelitian sebelumnya yang terkait mengenai penelitian ini, lalu pada sub-bagian ketiga (Sistem yang Dibangun) akan menjelaskan sistem dan teori yang digunakan pada penelitian ini. Pada sub-bagian keempat (Evaluasi) dan sub-bagian kelima (Kesimpulan) didapatkan Hasil, analisis dan kesimpulan yang diperoleh berdasarkan skenario pengujian yang telah di implementasikan pada penelitian ini.