

BAB I PENDAHULUAN

1.1 Latar Belakang

Pada pertengahan Mei 2022, UST dan Luna mengalami penurunan harga (*crash*) yang sangat signifikan dan merupakan *crash* terparah dalam sejarah kedua *cryptocurrency* tersebut[1]. Penurunan ini memicu perbincangan dan reaksi yang sangat intens dari masyarakat, terutama melalui media sosial seperti Twitter. Terlebih lagi banyaknya *tweet* dari orang-orang berpengaruh di bidang *cryptocurrency* mengenai opini mereka mengenai *crash* yang di alami kedua *cryptocurrency* tersebut, yang tentunya juga mempengaruhi opini dari para pengikutnya[2]. Dalam beberapa hari berturut-turut, topik ini menjadi *trending* pada platform twitter dengan rata-rata jumlah *tweet* harian pada bulan mei sebanyak 927 *tweet*.

Sentimen analisis adalah suatu teknik yang digunakan untuk mengekstraksi dan mengidentifikasi opini atau sentimen dari teks yang dianalisis[3]. Sentimen analisis pada Twitter dapat digunakan untuk mengetahui persepsi dan reaksi publik terhadap suatu peristiwa, khususnya dalam hal ini terkait *cryptocurrency crash* UST dan Luna. Namun, melakukan sentimen analisis pada data yang berlimpah dan beragam seperti pada Twitter memerlukan metode yang tepat dan efektif. Oleh karena itu pada penelitian ini akan dilakukan perbandingan algoritma klasifikasi dalam data sentimen dengan tiga metode klasifikasi yang dapat digunakan dalam melakukan penelitian mengenai sentimen analisis yaitu Metode *Multinomial Naive Bayes Classifier (MNB)*, *Support Vector Machine (SVM)* dan gabungan keduanya yaitu *Ensemble Learning*.

Multinomial Naïve Bayes Classifier(MNB) adalah metode *naïve bayes* yang biasa digunakan dalam klasifikasi data berupa teks, dimana fitur-fiturnya merupakan jumlah kemunculan suatu kata dalam sebuah dokumen[4]. *Support Vector Machine (SVM)* adalah metode *Machine Learning* yang digunakan untuk klasifikasi dan regresi, yang dapat efektif diterapkan untuk klasifikasi teks. SVM memiliki keunggulan dalam

menerapkan pemisahan linier pada input data non-linier dengan dimensi tinggi, yang diperoleh dengan menggunakan fungsi kernel yang diperlukan. Dengan memilih dan menerapkan fungsi kernel yang sesuai dengan karakteristik data, SVM dapat memberikan hasil yang akurat dalam mengklasifikasikan data teks[3]. Sedangkan *Ensemble learning* adalah teknik pembelajaran *Machine Learning* dimana sekumpulan model pembelajaran individu digabungkan sehingga model komponen saling mengimbangi kelemahan satu sama lain. Tujuan dari *ensemble learning* adalah untuk meningkatkan akurasi dan stabilitas dari model pembelajaran secara keseluruhan[5]. Atau merupakan gabungan dari beberapa model *machine learning*.

Penelitian ini bertujuan untuk Mendapatkan perbandingan dari ketiga metode di atas dalam melakukan sentimen analisis terhadap *cryptocurrency crash* Luna dan UST yang terdapat pada *Blockchain Terra*. Serta untuk memprediksi tingkat akurasi, *recall*, *F1-score*, *AUC (Area Under Curve)* dari kurva ROC dan *confusion matrix* yang akan menyimpulkan *True Positive Rate (TPR)* pada tiga model algoritma yang dipilih untuk melakukan klasifikasi sentimen. Hasil dari penelitian ini dapat menjadi acuan dalam pemilihan algoritma yang tepat dalam penelitian-penelitian selanjutnya.

1.2 Rumusan Masalah

Penelitian mengenai sentimen analisis umumnya hanya menggunakan satu model klasifikasi, dalam penelitian ini peneliti melakukan perbandingan pada 3 model klasifikasi yaitu MNB, SVM dan *Ensemble Learning* (gabungan dari SVM dan MNB). Dengan membandingkan parameter seperti nilai akurasi, presisi, *recall*, *F1-Score* *AUC* dari kurva ROC dan dan hasil *confusion matrix* dari tiap model klasifikasi, maka dapat dirumuskan suatu masalah yaitu :

1. Bagaimana hasil pengujian pada setiap model klasifikasi yang digunakan dalam melakukan sentimen analisis terkait insiden *cryptocurrency crash* UST dan Luna?
2. Model klasifikasi manakah yang menunjukkan hasil klasifikasi terbaik?

1.3 Batasan Masalah

Berikut adalah batasan masalah penelitian yang akan dilakukan dalam penelitian ini :

1. Penelitian ini menggunakan dataset yang diambil dari platform Twitter dari tanggal 01-Mei-2022 sampai tanggal 09-Juli-2022 dengan menggunakan teknik scraping, sehingga hanya *tweet* yang diposting oleh akun publik yang akan digunakan dalam penelitian ini.
2. Penelitian ini hanya memfokuskan pada sentimen yang terdapat dalam *tweet* terkait *crash cryptocurrency* Terra (UST dan Luna).
3. Penelitian ini melakukan perbandingan pada metode *Multinomial Naive Bayes Classifier (MNB)*, *Support Vector Machine (SVM)*, dan *Ensemble Learning* untuk mengklasifikasikan sentimen dalam *tweet*.
4. Penelitian ini hanya menggunakan *tweet* yang dikumpulkan dari platform Twitter sebagai data yang akan dianalisis.
5. Penelitian ini tidak memperhatikan *tweet* yang tidak terkait dengan *Cryptocurrency Crash* Terra (UST dan Luna).
6. Penelitian ini tidak memperhitungkan *tweet* yang berisi kata-kata yang tidak jelas atau tidak bermakna.
7. Penelitian ini tidak memperhitungkan *tweet* yang berisi emoticon atau emoji dalam mengklasifikasikan sentimen.
8. Penelitian ini menggunakan bahasa pemrograman Python dan hasil akhirnya adalah sebuah perbandingan sistem model klasifikasi.

1.4 Tujuan Penelitian

Berikut adalah tujuan dari penelitian ini adalah untuk :

1. Menentukan algoritma klasifikasi yang paling efektif dalam mengklasifikasikan sentimen analisis pada Twitter terkait insiden *Cryptocurrency Crash* UST dan Luna.

2. Menganalisis performa algoritma klasifikasi MNB, SVM, dan *Eensemble Learning* dalam mengklasifikasikan sentimen analisis pada Twitter terkait insiden *Cryptocurrency Crash* UST dan Luna.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat untuk peneliti diantaranya:

1. Menyediakan informasi tentang sentimen yang terdapat dalam *tweet* terkait crash *cryptocurrency* UST dan Luna.
2. Hasil penelitian dapat menunjukkan kemampuan metode *Multinomial Naive Bayes Classifier(MNB)*, *Support Vector Machine(SVM)*, dan *Ensemble Learning* dalam mengklasifikasikan sentimen dalam *tweet* terkait crash *cryptocurrency* UST dan Luna, sehingga dapat dijadikan sebagai acuan dalam penerapan metode tersebut pada penelitian yang serupa di masa mendatang.
3. Menjadi bahan referensi bagi peneliti lain yang ingin meneliti topik yang sama atau topik yang terkait dengan sentimen analisis pada Twitter.
4. Menambah pengetahuan dan wawasan tentang algoritma klasifikasi dan pengaplikasiannya dalam mengklasifikasikan sentimen analisis pada Twitter.

1.6 Sistematika Penulisan

Sistematika penulisan berisi gambaran umum mengenai penelitian yang akan dilakukan, sehingga mempermudah dalam melakukan pemahaman alur penulisan dalam penelitian. Berikut adalah gambaran singkat dari keseluruhan isi penulisan dalam penelitian yang berbentuk skripsi ini :

Bab I Pendahuluan, bagian ini merupakan bab awal yang berisi mengenai latar belakang penelitian, masalah penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian serta sistematika penulisan.

Bab II Landasan Teori, bagian ini berisikan tinjauan kepustakaan dari penelitian-penelitian terkait (biasanya berisikan dari penelitian terdahulu) dan juga berisikan beberapa teori yang berkaitan dengan penelitian yang akan dilakukan.

Bab III Metodologi Penelitian, bagian ini berisikan mengenai keseluruhan gambaran dari objek penelitian, alur penelitian, alat dan bahan yang menunjang penelitian, prosedur penelitian, serta metode yang diterapkan pada penelitian.

Bab IV Hasil dan Pembahasan, bagian ini membahas mengenai implementasi dari metode yang digunakan serta analisis dari hasil penelitian yang sudah dilakukan.

Bab V Penutup, Bagian ini menjelaskan tahap terakhir yang dilakukan oleh peneliti, yang mencakup ringkasan dari seluruh uraian sebelumnya serta kesimpulan yang ditarik. Tahap ini juga menunjukkan kelemahan dari penelitian ini serta saran untuk pengembangan penelitian berikutnya.

Daftar Referensi, Bagian ini berisi daftar referensi yang berkaitan dengan penelitian ini, termasuk publikasi jurnal dan artikel situs yang dapat membantu proses penelitian.