

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam beberapa tahun terakhir, mata uang kripto telah menjadi subjek perhatian global dan menjadi salah satu aset investasi yang paling menarik. Fenomena ini muncul karena mata uang kripto seperti Bitcoin dan Ethereum menggunakan teknologi *blockchain* yang memiliki karakteristik terdesentralisasi dan *immutable*, hal ini menjadikan *blockchain* teknologi yang aman, transparan serta efisien[1] sehingga dapat digunakan untuk berbagai tujuan seperti aset digital/investasi, media pertukaran online, pembayaran, dan penggunaan non-moneter di internet[2]. Pertumbuhan pesat dalam adopsi mata uang kripto telah mendorong perkembangan *platform crypto exchange* yang inovatif.

Salah satu *crypto exchange* resmi yang telah mendapat pengakuan di Indonesia adalah “Ajaib Kripto” yang berdiri pada tahun 2022 dengan nomor tanda daftar 008/BAPPEBTI/CP-AK/04/2022 dari BAPPEBTI. *Platform* ini menawarkan berbagai fitur sebagai daya tarik sehingga banyak orang memilih menggunakannya, hal ini tercermin dari jumlah unduhan di *Google PlayStore* yang sudah mencapai 100 ribu kali dalam kurun waktu satu tahun sejak dirilis. Selaras dengan fakta yang ada, banyak pengguna yang memberikan ulasan untuk mendeskripsikan pengalaman mereka ketika menggunakan *platform* tersebut. Ulasan-ulasan ini dapat diolah serta dimanfaatkan untuk mendapatkan pemahaman atau wawasan yang lebih baik dengan metode pendekatan yang tepat.

Dalam konteks ini, *natural language processing* (NLP) dapat digunakan untuk memproses, mengolah serta menginterpretasikan data hingga dapat bermanfaat[3]. Terdapat beberapa bidang utama dalam NLP, yaitu pemahaman bahasa alami, pemrosesan bahasa alami, analisis sentimen, *summarization*, dan *question answering*[4]. Bidang-bidang ini saling terkait dan sering digunakan bersama-sama untuk menyelesaikan tugas-tugas yang lebih kompleks, namun dengan tujuan untuk memahami pandangan ataupun opini, Sentimen Analisis merupakan bidang NLP yang paling relevan.

Sentimen analisis merupakan pendekatan *machine learning* yang digunakan untuk mengukur dan memahami perasaan, pandangan, atau opini yang dinyatakan oleh individu atau kelompok dalam konteks tertentu[5]. Terkait layanan Ajaib Kripto, Sentimen analisis berfokus pada memahami bagaimana pandangan dan opini terhadap pengalaman yang dirasakan oleh pengguna yang dinyatakan dalam bentuk ulasan. Hal ini penting karena sentimen yang terkandung didalam ulasan dapat mempengaruhi keputusan calon pengguna untuk memilih serta dapat menjadi alasan bagi pengembang untuk merumuskan kebijakan atau pembaruan bagi platformnya. Metode seperti *K-Nearest Neighbor*, *Random forest*, *Naïve bayes* serta *Support vector machine* merupakan metode *machine learning* yang bisa digunakan untuk sentimen analisis.

K-Nearest Neighbor (KNN) memiliki kelebihan yaitu sederhana, mudah dipahami dan efisien [6]. Pada penelitian yang dilakukan oleh Habibi dkk[7], Sentimen analisis dilakukan dengan cara membandingkan metode KNN dengan NBC. Dari penelitian tersebut didapatkan bahwa KNN menghasilkan nilai akurasi sebesar 72.37% sedangkan NBC menghasilkan nilai akurasi sebesar 67.84%. Kemudian pada penelitian yang dilakukan Nasution dan Hayaty[8], Sentimen Analisis dilakukan dengan cara membandingkan SVM dan KNN. Hasil yang didapatkan dari penelitian tersebut adalah SVM mendapatkan nilai akurasi sebesar 88.76%, lebih tinggi daripada KNN akan tetapi membutuhkan waktu proses yang jauh lebih lama. Waktu proses pada KNN yaitu 0.0140 detik, sedangkan pada SVM yaitu 0.0230 detik. Selanjutnya pada penelitian yang dilakukan oleh Ditendra dkk[9], Dalam penelitian tersebut dilakukan perbandingan antara KNN, NBC dan *Decision Tree* untuk analisis sentimen. Hasil yang didapatkan adalah KNN dan NBC menghasilkan nilai akurasi yang sama yaitu 56,71% sedangkan *Decision Tree* mendapatkan nilai akurasi sebesar 49,90%.

Dari penelitian-penelitian tersebut, *K-nearest Neighbor* (KNN) memiliki kinerja yang cukup baik dibandingkan dengan metode yang lain. Namun di balik kelebihan yang ada, KNN masih memiliki kemungkinan untuk melakukan kesalahan klasifikasi yang disebabkan oleh beberapa kelemahan KNN yaitu sensitif terhadap fitur yang tidak relevan, sulitnya menentukan Nilai K yang optimal, sensitif terhadap data berderau maupun data pencilan[10]. salah satu cara

untuk mengurangi dampak dari kelemahan tersebut adalah menerapkan metode *Bootstrap Agregating*.

Bootstrap Agregating merupakan metode yang mampu meningkatkan ketahanan model terhadap variasi data dan meningkatkan kehandalan prediksi [10]. Pada penelitian yang dilakukan oleh Aini dkk [11], Dilakukan perbandingan antara *Support Vector Machine*, *K-Nearest Neighbor*, *Naïve Bayes Classifier* dan penerapan *Bagging*. Dari penelitian tersebut didapatkan bahwa kombinasi *Bagging* dengan ke 3 metode klasifikasi dengan penerapan *oversampling* maupun tanpa *oversampling* memiliki akurasi paling tinggi daripada semua *single classifier* yang digunakan. Kemudian pada penelitian yang dilakukan oleh Arisani dan Muljono [12], Pada penelitian ini dilakukan perbandingan antara KNN Dasar, KNN dengan Prapemrosesan dan KNN dengan *Bagging*. Hasil dari penelitian ini menunjukkan bahwa kombinasi KNN dan *Bagging* memiliki akurasi 95.68%, lebih unggul dari KNN+Pemrosesan yang mendapatkan akurasi 94.05%, KNN Dasar 85.41%. Berdasarkan pada penelitian-penelitian tersebut *Bootstrap Aggregation* dapat digunakan untuk meningkatkan akurasi dari metode klasifikasi.

Penelitian ini akan dilakukan dengan tujuan melakukan sentimen analisis pada ulasan *platform* bursa kripto bernama “Ajaib Kripto” yang ada di *Google Play Store* dengan pendekatan *machine learning* menggunakan kombinasi metode *K-Nearest Neighbor* dan *Bootstrap Aggregating*, di harapkan hasil penelitian ini dapat berguna bagi penelitian di masa yang akan datang khususnya di bidang sentimen analisis.

1.2 Rumusan Masalah

1.2.1 Permasalahan

Analisis sentimen pada ulasan *crypto exchange* “Ajaib Kripto” merupakan salah satu langkah penting untuk menilai kualitas layanan dari *exchange* tersebut. Algoritma KNN dapat melakukan klasifikasi sentimen pada ulasan dengan performa prediksi yang cukup baik, namun masih terdapat kemungkinan terjadinya kesalahan klasifikasi, salah satu cara untuk mengurangi dampak akibat permasalahan tersebut yaitu dengan menerapkan *Bootstrap Aggregating* yang dapat menggabungkan prediksi dari beberapa model untuk mendapatkan akurasi prediksi yang lebih baik.

1.2.2 Solusi Permasalahan

Penelitian ini mengusulkan kombinasi metode *K-Nearest Neighbor* dan *Bootstrap Aggregating* untuk melakukan analisis sentimen pada ulasan pengguna “Ajaib Kripto” di *Google Play Store*.

1.2.3 Pertanyaan Penelitian

1. Berapa performa metode *K-Nearest Neighbor* dengan nilai hyperparameter *K* terbaik pada analisis sentimen ulasan pengguna “Ajaib Kripto” ?
2. Berapa performa metode *K-Nearest Neighbor* + *Bagging* nilai hyperparameter *K* terbaik pada analisis sentimen ulasan pengguna “Ajaib Kripto” ?

1.3 Tujuan Penelitian

1. Untuk mengetahui Nilai *K* optimal metode *K-Nearest Neighbor* pada analisis sentimen ulasan pengguna “Ajaib Kripto”.
2. Untuk mengetahui performametode *K-Nearest Neighbor* dengan penerapan *Bootstrap Aggregating* pada analisis sentimen ulasan pengguna “Ajaib Kripto”.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah untuk mengetahui pengaruh *Bootstrap Aggregating* terhadap performa algoritma KNN dan menciptakan suatu model *machine learning* yang dapat melakukan sentimen analisis pada ulasan atau opini berdasar sentimen yang terkandung. Hasil penelitian ini juga akan menunjukkan bagaimana sentimen pengguna terhadap *platform* Ajaib Kripto serta diharapkan penelitian ini dapat berkontribusi bagi penelitian di masa yang akan datang, khususnya untuk bidang analisis sentimen.

1.5 Batasan Masalah

Terdapat beberapa batasan dalam penelitian ini agar lingkup penelitian tetap terfokus dan hasil yang diperoleh dapat diinterpretasikan dengan baik. Berikut adalah beberapa batasan yang diterapkan dalam penelitian ini:

1. Dataset yang di gunakan berasal dari ulasan *Google Play Store*
2. Pelabelan berdasarkan rating pada ulasan
3. Bahasa pemrograman yang digunakan adalah *python*

4. Label sentimen di bagi menjadi 2, yaitu positif dan negatif
5. Matriks jarak yang di gunakan adalah *Euclidean distance*

1.6 Sistematika Penulisan

Sistematika penulisan laporan yang digunakan terdiri dari beberapa bagian seperti berikut :

BAB 1 : PENDAHULUAN

Pada bab ini memuat latar belakang masalah, perumusan masalah, tujuan dan manfaat, batasan penelitian, dan sistematika penulisan.

BAB 2: TINJAUAN PUSTAKA

Pada bab ini, akan diuraikan landasan teori dan penelitian terdahulu yang relevan dengan penelitian yang diajukan.

BAB 3: METODE USULAN

Pada bab ini, akan diuraikan metode yang diusulkan, termasuk arsitektur sistem, dataset dan tahapan-tahapan penelitian yang akan dilakukan.

BAB 4: HASIL DAN PEMBAHASAN

Pada bab ini, akan di uraikan hasil dan pembahasan dari penelitian yang telah di lakukan.

BAB 5: KESIMPULAN DAN SARAN

Pada bab ini, akan di lakukan penarikan kesimpulan serta saran dari penelitian yang telah di lakukan.