

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan data yang signifikan di dunia digital modern, perkembangan data telah bertransformasi menjadi sumber daya strategis yang mendukung berbagai sektor kehidupan. Seiring meningkatnya volume data yang dihasilkan secara masif, khususnya melalui platform media sosial seperti X, pengelolaan yang tidak tepat berisiko menjadikan data tersebut tidak bernilai dan terabaikan. Salah satu bentuk data tersebut adalah *tweet*, yang terus bertambah setiap hari dan menyimpan potensi informasi yang signifikan. Untuk mengoptimalkan nilai informasi dari himpunan data besar seperti ini, berkembanglah suatu pendekatan analitis yang sangat populer, yakni *Text Mining*, yang memungkinkan ekstraksi informasi relevan dari data tekstual secara sistematis dan efisien.(Parlika et al., 2020), hal tersebut menjadi semakin penting seiring dengan pertumbuhan data yang eksponensial, sehingga penambangan teks berfungsi sebagai mekanisme utama dalam mengubah data mentah menjadi pengetahuan yang terstruktur.

Dalam konteks perkembangan data digital, metode penambangan teks telah menjadi sarana yang semakin dibutuhkan dalam menggali pengetahuan dari data masif yang tidak terorganisir. Salah satu penerapan utama dari teknik ini adalah analisis sentimen, yang memungkinkan pemetaan opini publik terhadap berbagai isu, termasuk investasi dalam aset digital seperti *cryptocurrency*. Dengan meningkatnya minat terhadap mata uang kripto, analisis sentimen berbasis strategi *mining text* untuk menelusuri bagaimana sudut pandang masyarakat terhadap volatilitas pasar, aturan, serta potensi peluang ancaman dalam investasi dan keuntungan dalam investasi kripto. (Zhuo et al., 2024), hal ini menunjukkan bahwa analisis sentimen tidak hanya bermanfaat bagi pengambil kebijakan, tetapi juga relevan dalam menilai risiko dan peluang yang melekat pada aset digital. dari hasil riset terbaru memperlihatkan bahwa pendekatan ini dapat memberikan wawasan yang lebih akurat dalam mengidentifikasi pola sentimen yang berkembang di media sosial dan berita terkait *cryptocurrency*, sehingga

membantu investor dan regulator dalam mengambil keputusan yang lebih berbasis data.

Analisis sentiment ini sangat berperan penting dalam membantu masyarakat memahami berbagai pandangan terhadap suatu topik. contoh dalam konteks koin kripto, meskipun tingkat fluktuasi nilai yang tinggi, risiko kejahatan siber, serta potensi praktik pencucian uang akibat sifat anonimitasnya menjadi perhatian utama. Penelitian yang saya lakukan bertujuan untuk mengkaji perspektif masyarakat terhadap investasi kripto dengan mengelompokkan sentimen ke dalam dua kategori : positif dan negatif.(Chandra Waskita & Santika, 2022).

Dari Berbagai penelitian sebelumnya telah mengeksplorasi analisis sentiment menggunakan beragam pendekatan. berdasarkan studi terdahulu, penelitian ini berfokus pada analisis komentar terkait investasi kripto dengan strategi klasifikasi menggunakan K-NN, Tujuan utama penelitian adalah mendalami bagaimana masyarakat menanggapi terhadap investasi kripto di platform media sosial, khususnya X. kajian berupaya untuk menghadirkan solusi dalam memahami opini masyarakat terhadap investasi kripto.

Kajian ini memberikan memberikan nilai penting untuk pengembangan studi investasi kripto melalui pendekatan yang menitikberatkan pada analisis sentiment terhadap investasi crypto.yang bertujuan untuk memahami bagaimana pandangan masyarakat Indonesia terhadap investasi dalam aset digital seperti *cryptocurrency*. Dalam upaya mencapai tujuan tersebut, menjalankan strategi model klasifikasi berbasis K-NN pada kajjian ini, adalah salah satu metode dalam *machine learning* yang membantu identifikasi pola dan kecendrungan sentimen dengan memanfaatkan informasi yang bersumber dari media social X.

Analisis sentiment ini juga digunakan untuk mengevaluasi performa algoritma KNN dalam mengklasifikasikan komentar masyarakat terkait investasi kripto di Indonesia. Dengan memahami pola sentiment ini, masyarakat dapat memperoleh wawasan lebih mendalam mengenai tren investasi kripto, memantau perkembangannya di Indonesia, serta menggunakan temuan tersebut dapat dijadikan rujukan untuk menimbang penentuan langkah yang diambil. untuk melakukan analisis sentimen, beberapa aspek penting harus diperhatikan,

termasuk pemilihan dataset yang akan di gabung acuan dalam proses pengolahan data.

1.2 Rumusan Masalah

1. Bagaimana sentiment masyarakat Indonesia terhadap investasi *cryptocurrency* berdasarkan data komentar di X (Twitter)?
2. Bagaimana prperforma kinerja model dalam konteks overfitting di algoritma K-Nearest Neighbor (KNN)?
3. Bagaimana performa algoritma KNN dalam mengklasifikasikan sentiment positif dan negatif berdasarkan data yang diperoleh?

1.3 Tujuan Penelitian

1. Mengidentifikasi sentiment masyarakat Indonesia terhadap investasi *cryptocurrency* melalui data komentar di platform X.
2. Mengidentifikasi performa kinerja model dalam konteks overfitting di algoritma K-Nearest Neighbor (KNN).
3. Mengevaluasi performa algoritma K-NN dalam klasifikasi sentiment berdasarkan metrik akurasi, presisi, dan recall.

1.4 Manfaat Penelitian

1. Manfaat dari kajian ini yaitu mengetahui performa algoritma K – Nearest Neighbor dalam melakukan klasifikasi berdasarkan komentar X yang ada di indonesia.
2. Menjadi referensi bagi peneliti yang akan datang, yang Dari penelitian ini diharapkan agar dapat menambah pengetahuan mengenai investasi crypto

1.5 Batasan Penelitian

1. Model K – Nearest Neighbor yang digunakan bergantung pada kualitas dataset dan preprocessing, sehingga bisa mengalami keterbatasan dalam menangani kontek kompleks.
2. Data yang dianalisis hanya berasal dari X, tidak mencakup platform lain seperti Facebook, Instagram, atau forum diskusi.

3. Tweet yang mengandung singkatan atau istilah slang hanya diproses menggunakan metode dasar tanpa penyesuaian dengan kamus khusus bahasa Indonesia.
4. penelitian ini hanya menggunakan tools dari library node.js yaitu [tweet-harvest@2.6.1](#) sebagai Teknik pengumpulan data
5. Penelitian ini dibatasi pada penggunaan Google Colab dengan bahasa pemrograman Python sebagai media pengolahan dan analisis data.

