

BAB I

PENDAHULUAN

1.1 Latar Belakang

Teknologi informasi dan internet telah memberikan banyak manfaat dalam berbagai kehidupan masyarakat, ekonomi, perdagangan ataupun investasi. *Bitcoin* dan kripto adalah salah satu aset atau investasi berbentuk digital yang bukan diakui sebagai alat pembayaran yang sah. Seiring meningkatnya minat masyarakat terhadap aset digital, keberadaan *platform* perdagangan menjadi sangat penting untuk memfasilitasi akses yang aman dan mudah. Peran *platform* perdagangan menjadi sangat penting untuk mendukung akses masyarakat terhadap aset digital, termasuk *bitcoin* dan kripto. Pertumbuhan jumlah pengguna dan transaksi ini tidak hanya menunjukkan meningkatnya popularitas aset kripto di Indonesia, tetapi juga menegaskan pentingnya pengalaman pengguna dalam menggunakan aplikasi tersebut. Salah satu cara untuk memahami pengalaman dan kepuasan pengguna adalah melalui ulasan yang diberikan pada *platform* seperti *Google Play Store* (Sunarko, Wasino, dan Sutrisno 2023)

Aplikasi *Indodax* merupakan salah satu *platform* perdagangan aset kripto terbesar di Indonesia. Pengguna aplikasi *Indodax* telah mencapai 7,4 juta member terdaftar (Jual Beli Bitcoin dan Trading Kripto Indonesia - INDODAX 2025). *Indodax* telah mendapatkan berbagai ulasan berupa saran, kritik maupun keluhan pada *platform Google Play Store*. Hal itu menjadi sangat bermanfaat untuk pengguna lain yang akan menggunakan aplikasi tersebut (Anton dkk. 2023).

Sentimen yang terkandung dalam ulasan-ulasan ini dapat mencerminkan sejauh mana aplikasi ini memenuhi ekspektasi, memberikan layanan yang andal dan memperbaiki kendala teknis yang ada serta analisis sentimen dapat menghasilkan kesimpulan dari data yang akan dianalisis sehingga dapat memenuhi kebutuhan pengguna (Aida Sapitri dan Fikry 2023). Untuk membantu proses analisis sentimen adalah dengan memanfaatkan teknologi *machine learning* yang dapat membantu mengolah data ulasan secara otomatis dan efektif.

Dalam menganalisis sentimen pengguna aplikasi *Indodax*, berbagai metode *Machine Learning* digunakan untuk membantu proses analisis. *Support Vector Machine*, *Random Forest*, *Extreme Gradient Boosting (XGBoost)*. *Support Vector Machine* digunakan untuk menemukan batas pemisah terbaik diantara berbagai kelas sentimen (Oktavia dan Ramadhan 2023). Penelitian sebelumnya mengenai penggunaan *Support Vector Machine* yang dilakukan oleh (Wahyudi dan Kusumawardana 2021) pada analisis sentimen ulasan aplikasi Grab, dengan hasil akurasi sebesar 85,54%. *Random Forest* membantu mengurangi *overfitting* dengan menggabungkan beberapa pohon keputusan (Mardiyati dkk. 2023). Penelitian sebelumnya mengenai penggunaan *Random Forest* telah dilakukan oleh (Yasir dkk. 2024) pada analisis sentimen terhadap kontroversi Fatwa MUI, dengan hasil akurasi sebesar 67%. *Extreme Gradient Boosting (XGBoost)* pendekatan pada data sentimen dengan memanfaatkan teknik *boosting* untuk meningkatkan akurasi model (Kurniawan dan Indahyanti 2024). Penelitian sebelumnya mengenai penggunaan *Extreme Gradient Boosting (XGBoost)* yang dilakukan oleh (Jayanti

dkk. 2024) pada analisis sentimen ulasan aplikasi belanja online, dengan hasil akurasi sebesar 71%.

Penelitian sebelumnya yang dilakukan oleh (Erkamim dkk. 2023) memiliki kekurangan khususnya pada fitur yang redundan. Solusi untuk mengatasi permasalahan yang terjadi seperti fitur yang redundan adalah dengan menambahkan *Feature Selection Chi-Square*. *Feature Selection Chi-Square* adalah pemilihan fitur yang diawasi sehingga dapat menghilangkan banyak fitur untuk memungkinkan menaikkan tingkat akurasi (Erkamim dkk. 2023). Fitur-fitur yang tidak relevan atau kurang informatif dapat dihilangkan, sehingga meningkatkan efisiensi dan akurasi model dalam analisis sentimen atau klasifikasi lainnya (Ernayanti dkk. 2023)

Berdasarkan pada penelitian terkait sebelumnya, terdapat sebuah peluang penelitian yaitu melakukan perbandingan dengan menerapkan *Feature Selection Chi-Square* untuk membandingkan *Support Vector Machine*, *Random Forest* dan *XGBoost* agar dapat menentukan salah satu algoritma yang mendapatkan tingkat akurasi paling tinggi. Penelitian ini diharapkan memperoleh performa algoritma yang optimal dengan mempertimbangkan jumlah fitur yang paling berkontribusi pada akurasi, sehingga proses klasifikasi dapat dilakukan secara lebih efisien dan efektif.

1.2 Rumusan Masalah

Berdasarkan penjelasan dalam latar belakang sebelumnya, maka permasalahan yang akan dibahas adalah sebagai berikut:

1. Bagaimana performa masing-masing algoritma *Support Vector Machine* (SVM), *Random Forest*, dan *XGBoost* dalam mengklasifikasikan sentimen ulasan pengguna aplikasi *Indodax*, ditinjau dari akurasi dan efisiensi komputasi dengan penerapan *Feature Selection Chi-Square*?
2. Algoritma manakah yang menunjukkan performa klasifikasi sentimen paling optimal setelah penerapan metode *Feature Selection Chi-Square* dalam mengatasi masalah fitur redundan pada data ulasan pengguna?

1.3 Tujuan Penelitian

Tujuan penelitian ini dirumuskan untuk memberikan arah yang jelas dalam mencapai hasil yang diharapkan, terdapat beberapa tujuan dari penelitian ini yang ingin dicapai, yaitu:

1. Membandingkan performa algoritma *Support Vector Machine* (SVM), *Random Forest*, dan *XGBoost* dalam klasifikasi sentimen ulasan aplikasi *Indodax* berdasarkan akurasi prediksi dan efisiensi komputasi, dengan menerapkan metode *Feature Selection Chi-Square*.
2. Mengevaluasi dan menentukan model klasifikasi sentimen yang paling optimal, serta menggunakan seleksi fitur Chi-Square untuk meminimalkan redundansi fitur yang memungkinkan menaikkan akurasi pada model.

1.4 Manfaat Penelitian

Adapun manfaat dari dilakukannya penelitian ini, diantaranya:

1. Memberikan wawasan dalam perbandingan beberapa metode klasifikasi sentimen dengan menggunakan *Feature Selection Chi-Square*.

2. Meningkatkan efisiensi proses klasifikasi dengan memilih fitur relevan yang dapat mengurangi kompleksitas model.
3. Hasil penelitian ini dapat membantu mengembangkan metode analisis sentimen yang lebih efektif dan efisien untuk berbagai aplikasi lainnya.

1.5 Batasan Masalah

Penelitian menetapkan batasan-batasan yang jelas agar menjadi lebih spesifik. Batasan masalah ini akan menentukan ruang lingkup penelitian serta membantu dalam membatasi variabel yang diteliti. Adapun batasan-batasan masalah yang ditetapkan dalam penelitian ini adalah sebagai berikut:

1. Pengumpulan data ulasan *Indodax* hanya diambil dari *Google Play Store* secara *real-time*.
2. Metode yang digunakan *Support Vector Machine*, *Random Forest* dan *XGBoost*.
3. Tidak menambahkan *hyperparameter tuning* pada model *Support Vector Machine*, *Random Forest* dan *XGBoost*.