BAB I
PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital membuat aktivitas belanja *online* semakin mudah dan praktis. Dukungan akses internet yang luas, metode pembayaran digital yang beragam, serta promosi yang menarik telah mendorong pertumbuhan *e-commerce* secara pesat [1]. Namun, pertumbuhan ini juga membawa tantangan berupa maraknya kasus penipuan, seperti transaksi fiktif [2], pencurian identitas [3], penggunaan kartu kredit curian [4], hingga manipulasi metode pembayaran [1]. Kondisi ini tidak hanya merugikan secara finansial, tetapi juga menurunkan kepercayaan konsumen terhadap keamanan platform *e-commerce* [5].

Seiring dengan meningkatnya resiko penipuan digital, pelaku industri dan pengguna menjadi semakin khawatir [5]. Sistem keamanan yang ada seringkali belum mampu mendeteksi pola kecurangan secara cepat dan tepat, terutama ketika volume transaksi meningkat pesat setiap harinya [2]. Perkembangan *e-commerce* yang sangat pesat telah menyebabkan lonjakan jumlah transaksi daring yang diikuti oleh meningkatnya aktivitas penipuan [6]. Oleh karena itu, dibutuhkan pendekatan berbasis teknologi yang lebih adaptif dan cerdas untuk mengidentifikasi penipuan sejak dini serta meminimalkan kerugian yang ditimbulkan [7].

Machine learning (ML) menawarkan model yang mampu mengenali pola dari data besar secara akurat dan efisien, sehingga menjadi alat analisis cerdas yang banyak digunakan untuk meningkatkan akurasi deteksi dan efisiensi operasional [8]. Teknologi ini memungkinkan analisis data historis untuk mendeteksi aktivitas mencurigakan secara *real-time* [6]. Dengan memanfaatkan data historis transaksi, system dapat belajar mengenali tidak biasa yang menandakan penipuan, sehingga deteksi dapat dilakukan secara cepat sebelum kerugian terjadi [9]. Beberapa penelitian menunjukkan algoritma *ML* meningkatkan akurasi deteksi penipuan digital [3]. Namun, sebagian penelitian masih berfokus pada akurasi tanpa

memperhatikan efisiensi waktu pemrosesan. Padahal, efisiensi dan kemudahan interpretasi penting untuk mendukung pengambilan keputusan cepat dan tepat.

Penelitian ini menggunakan algoritma *Random Forest* dan *Naive Bayes* karena keduanya memiliki pendekatan berbeda namun sama-sama efektif dalam mendeteksi pola penipuan digital pada transaksi *e-commerce* [10]. *Random Forest* dikenal memiliki akurasi tinggi dan mampu menilai fitur yang berpengaruh terhadap hasil klasifikasi [11], sedangkan *Naive Bayes* unggul dalam efisiensi komputasi dan kecepatan pemrosesan data [12]. Dengan membandingkan kedua algoritma ini, penelitian bertujuan menentukan metode yang paling optimal dalam mendeteksi penipuan digital secara akurat dan efisien.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang yang telah dijelaskan sebelumnya, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana kinerja algoritma *Random Forest* dan *Naive Bayes* dalam mendeteksi transaksi penipuan pada data *e-commerce* berdasarkan evaluasi metrik?
2. Metode mana yang paling optimal antara algoritma *Random Forest* dan *Naive Bayes* dalam mendeteksi transaksi penipuan pada *e-commerce* berdasarkan hasil evaluasi performa model?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah dan fokus pada tujuan yang ingin dicapai, maka batasan masalah yang diterapkan adalah sebagai berikut:

1. Penelitian ini berfokus pada perbandingan kinerja dua algoritma klasifikasi, yaitu *Random Forest* dan *Naive Bayes*, dalam mendeteksi transaksi penipuan pada database *e-commerce*.
2. Data yang digunakan berasal dari dataset transaksi *e-commerce* publik, yang memuat informasi seperti waktu transaksi, nilai pembelian, sumber akses, perangkat, jenis kelamin pengguna, serta status transaksi *(fraud* atau

non-fraud).

3. Proses penelitian meliputi *preprocessing* data, penanganan ketidakseimbangan kelas menggunakan *RUS*, pelatihan model, dan evaluasi performa algoritma berdasarkan metrik *Accuracy, Precision, Recall, F1-score, ROC-AUC, False Positive Rate (FPR), dan False Negative Rate (FNR).*

4. Analisis difokuskan pada performa dan efisiensi komputasi algoritma dalam skenario pengujian berbasis data, tanpa membahas penerapan model secara langsung pada sistem *e-commerce* nyata.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Menganalisis dan membandingkan kinerja algoritma *Random Forest* dan *Naive Bayes* dalam mendeteksi transaksi penipuan pada data *e-commerce* berdasarkan metrik evaluasi seperti *Accuracy, Precision, Recall, F1-score, ROC-AUC, False Positive Rate (FPR), dan False Negative Rate (FNR).*

2. Menentukan algoritma yang paling optimal antara *Random Forest* dan *Naive Bayes* dalam mendeteksi transaksi penipuan pada *e-commerce* berdasarkan hasil evaluasi performa model, meliputi tingkat ketepatan prediksi, efisiensi waktu pemrosesan, dan efektivitas komputasi.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat Teoritis

    1. Memberikan kontribusi terhadap pengembangan ilmu di bidang data mining dan machine learning, khususnya dalam penerapan algoritma klasifikasi untuk deteksi penipuan digital pada *e-commerce*.

    2. Menjadi referensi akademik bagi penelitian selanjutnya yang ingin melakukan studi perbandingan efektivitas algoritma

menggunakan metrik evaluasi seperti *Accuracy, Precision, Recall, F1-score, ROC-AUC, False Positive Rate (FPR), dan False Negative Rate (FNR).*

3. Menambah wawasan mengenai perbandingan performa algoritma *Random Forest* dan *Naive Bayes* dalam konteks pendeteksian transaksi penipuan.

2. Manfaat Praktis

1. Menjadi acuan bagi pengembang sistem keamanan *e-commerce* dalam menentukan algoritma yang optimal untuk mendeteksi transaksi penipuan secara cepat dan efisien.

2. Membantu pelaku industri digital meningkatkan keandalan sistem keamanan serta mengurangi potensi kerugian akibat transaksi penipuan.

3. Mendukung peningkatan kepercayaan konsumen terhadap keamanan transaksi pada platform *e-commerce.*

1.6 Sistematika Penulisan

Sistematika penulisan skripsi ini disusun agar memudahkan pembaca dalam memahami isi penelitian secara terstruktur. Adapun susunan skripsi ini terdiri dari lima bab sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi uraian mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan skripsi. Bab ini memberikan gambaran umum mengenai alasan dilakukannya penelitian serta arah penelitian yang akan dilakukan.

BAB II TINJAUAN PUSTAKA

Bab ini memuat studi literatur yang berkaitan dengan topik penelitian, termasuk hasil-hasil penelitian terdahulu mengenai deteksi penipuan *e-commerce* menggunakan algoritma *machine learning*. Selain itu, disajikan pula dasar teori yang mendukung penelitian, seperti teori mengenai algoritma *Random Forest*, *Naïve Bayes*, evaluasi model klasifikasi, dan konsep dasar fraud detection.

BAB III METODE PENELITIAN

Bab ini menjelaskan tentang objek penelitian, alur penelitian, serta alat dan bahan yang digunakan. Di dalamnya juga dijelaskan tahapan-tahapan penelitian seperti *preprocessing* data, pembagian dataset, penerapan algoritma, hingga metode evaluasi performa model. Bab ini bertujuan menggambarkan langkah-langkah sistematis yang dilakukan peneliti dalam melaksanakan penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi hasil implementasi dan pengujian algoritma Random Forest serta Naive Bayes pada dataset e-commerce. Disajikan pula analisis hasil perbandingan performa kedua algoritma berdasarkan metrik evaluasi seperti *Accuracy*, *Precision*, *Recall*, *F1-score*, *ROC-AUC*, *False Positive Rate (FPR)*, dan *False Negative Rate (FNR)*. Pembahasan dilakukan untuk menjawab rumusan masalah dan tujuan penelitian.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan serta saran untuk pengembangan penelitian selanjutnya. Kesimpulan disusun berdasarkan hasil analisis dan pembahasan yang telah dijabarkan pada bab sebelumnya.