

BAB I

PENDAHULUAN

1.1 Latar Belakang

Transaksi jual beli elektronik, atau terkenal dengan sebutan *e-commerce*, pada era digital sekarang ini telah menjadi salah satu cara yang paling diminati konsumen untuk berbagi produk maupun jasa [1]. Namun, dengan meningkatnya kompetisi pada pasar *e-commerce*, perusahaan perlu memahami perilaku konsumen dengan lebih baik untuk meningkatkan penjualan dan loyalitas pelanggan. Melalui analisis perilaku konsumen dapat membantu perusahaan untuk mengidentifikasi pola dan kecenderungan untuk pengembangan teknik strategi pemasaran serta meningkatkan pengalaman pelanggan [2].

Salah satu cara untuk analisis perilaku konsumen yaitu dengan memanfaatkan algoritma pembelajaran mesin (*machine learning*) untuk mengidentifikasi pola dan kecenderungan pada data konsumen lalu memprediksinya. Selama beberapa tahun terakhir, algoritma *Random Forest* dan *eXtreme Gradient Boosting (XGBoost)* tengah populer untuk menganalisis perilaku konsumen karena memiliki banyak keunggulan. Hal ini karena kedua algoritma tersebut lebih *robust* terhadap *outliers* (data pencilan), sehingga memungkinkan efisiensi waktu ketika proses komputasi dan mampu memberikan akurasi yang tepat walaupun menangani *dataset* besar dan kompleks [3]. Namun, ditengah kepopulerannya, banyak perdebatan di dunia penelitian untuk menentukan mana algoritma yang kompatibel pada permasalahan analisis perilaku konsumen di antara kedua model tersebut.

Beberapa penelitian menunjukkan bahwa *Random Forest* memiliki kinerja lebih optimal dalam menangani distribusi data tidak seimbang, sedangkan *XGBoost* memiliki kinerja lebih optimal dalam menangani data yang besar dan kompleks [4]. Hal ini dibuktikan melalui analisis perilaku konsumen pada platform *social media* dengan memanfaatkan algoritma *Random Forest*, di mana penelitian ini mampu mencapai nilai akurasi sebesar 97.89% [5]. Sebaliknya, pada analisis perilaku konsumen berbasis *website*, algoritma *XGBoost* lebih unggul karena memperoleh nilai F1 yang tinggi, yaitu sebesar 92% [6].

Oleh sebab itu, penelitian ini akan melakukan studi komparatif mengenai algoritma *Random Forest* dan *XGBoost* untuk analisis perilaku konsumen pada industri *e-commerce*. Penelitian memanfaatkan *dataset e-commerce* berukuran besar dan kompleks untuk membandingkan kinerja kedua algoritma tersebut.

1.2 Rumusan Masalah

Berikut rumusan masalah pada penelitian ini, antara lain:

1. Bagaimana kinerja algoritma *Random Forest* dan *XGBoost* dalam analisis perilaku konsumen *e-commerce*?
2. Mana algoritma yang lebih baik dalam menangani data besar yang tidak seimbang dan kompleks dalam analisis perilaku konsumen *e-commerce*?

1.3 Batasan Masalah

Berikut batasan masalah yang digunakan dalam penelitian ini, antara lain:

1. Penelitian ini hanya akan menggunakan *dataset e-commerce* yang besar dan kompleks.
2. Penelitian ini hanya akan menggunakan algoritma *Random Forest* dan *XGBoost* untuk menganalisis perilaku konsumen, khususnya loyalitas pelanggan.
3. Penelitian ini tidak akan menggunakan *model tuning* untuk meningkatkan kinerja algoritma.
4. Penelitian ini dijalankan dengan memanfaatkan versi *free tier* pada mesin *google colaboratory*.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk melakukan perbandingan antara dua algoritma *machine learning* untuk menganalisis perilaku konsumen, yang merupakan topik dalam ilmu komputer dan *data science*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan metode analisis perilaku konsumen *e-commerce* yang lebih baik. Hasil penelitian ini dapat digunakan oleh perusahaan *e-commerce* untuk meningkatkan strategi

pemasaran dan meningkatkan pengalaman pelanggan. Secara teknis, penelitian ini dapat membantu untuk meningkatkan kemampuan algoritma *machine learning* untuk menganalisis perilaku konsumen *e-commerce*.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN

Bab ini mengkaji perihal latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini memberikan pemahaman awal tentang masalah yang akan diteliti berdasarkan kajian pustaka yang relevan pada penelitian sebelumnya. Kajian ini mencakup studi literatur dan dasar teori penelitian.

BAB III METODE PENELITIAN

Bab ini memberikan gambaran umum objek penelitian, penjelasan tentang proses alur penelitian serta alat dan bahan yang digunakan untuk menunjang penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas mengenai hasil uji coba terhadap model yang digunakan, kemudian analisis hasil penelitian tersebut dituangkan dalam bentuk laporan penelitian.

BAB V PENUTUP

Bab ini menyampaikan kesimpulan dari keseluruhan analisis penelitian dan memberikan saran penelitian kedepannya yang berkaitan dengan kekurangan penelitian ini sehingga dapat mendorong pengembangan penelitian lebih lanjut.