

BAB I

PENDAHULUAN

1.1 Latar Belakang

IndiHome, produk layanan PT Telekomunikasi Indonesia (Telkom) yang menawarkan paket komunikasi dan data terpadu (telepon rumah, internet, dan televisi interaktif), telah menjadi pilihan populer di Indonesia (Sendari, 2021). Namun, kepuasan pelanggan terhadap IndiHome dipengaruhi oleh berbagai faktor, termasuk kualitas layanan, kecepatan dan kestabilan internet, aksesibilitas, dan biaya.

Umpan balik pelanggan mengenai IndiHome seringkali disampaikan melalui media sosial, terutama Twitter (kini X) (Jemadu, 2022). Platform ini memungkinkan pengguna untuk berbagi opini dan pengalaman mereka secara singkat dan langsung (Gil & Fisher, 2023). Oleh karena itu, analisis sentimen terhadap komentar-komentar di Twitter dapat memberikan wawasan berharga bagi Telkom untuk meningkatkan layanannya.

Analisis sentimen merupakan pendekatan efektif untuk memahami opini publik secara otomatis. Metode ini mampu mengolah teks dalam jumlah besar dengan efisien dan akurat, bahkan melebihi kemampuan manusia dalam beberapa kasus. Salah satu teknik analisis sentimen yang populer adalah penggunaan

machine learning, di mana model dilatih menggunakan dataset berlabel untuk mengklasifikasikan sentimen menjadi *positive*, *negative*, atau *neutral*.

Algoritma XGBoost, salah satu metode *machine learning* yang telah banyak digunakan, menawarkan kinerja yang unggul dalam berbagai skenario. XGBoost merupakan algoritma *tree ensemble* yang terdiri dari kumpulan *classification and regression trees* (CART). Keunggulan utama XGBoost terletak pada skalabilitas dan efisiensinya yang tinggi dimana XGBoost mampu menangani dataset berjumlah besar secara efektif dan cepat, berkat optimasi dari algoritma sebelumnya dan inovasi seperti *novel tree learning algorithm* untuk menangani data sparse (Chen & Guestrin, 2016).

Selain itu, model algoritma XGBoost juga dikenal memiliki beberapa keunggulan lainnya seperti tingkat akurasi yang tinggi, fleksibilitas yang luas dengan mendukung berbagai tipe dan tujuan data, teknik regularisasi untuk menghindari overfitting, dan interpretabilitas untuk memberikan skor tingkat kepentingan fitur yang dapat membantu pengguna memahami fitur mana yang paling penting untuk membuat prediksi (Vidhya, 2024).

Berdasarkan pertimbangan tersebut terutama pada keunggulan utama algoritma XGBoost seperti skalabilitas dan efisiensi, maka penelitian ini dilaksanakan dengan tujuan untuk menganalisis sentimen terhadap layanan IndiHome berdasarkan pendapat pelanggan yang diungkapkan di sosial media menggunakan metode *eXtreme Gradient Boosting* (XGBoost).

1.2 Rumusan Masalah

Berdasarkan latar belakang sebelumnya maka rumusan masalah dari penelitian sebagai berikut:

1. Bagaimana performa model klasifikasi sentimen menggunakan algoritma XGBoost dalam memprediksi sentimen *tweet* dengan kata kunci "indihome"?
2. Bagaimana karakteristik sentimen yang terungkap dari hasil klasifikasi *tweet* terkait IndiHome menggunakan XGBoost?

1.3 Batasan Masalah

Batasan masalah dalam melakukan proses dari penelitian ini yaitu:

1. Penelitian ini hanya akan menganalisis sentimen dari *tweet* dengan *keyword* 'indihome' yang dikumpulkan dari sosial media Twitter (X).
2. Periode pengambilan data *tweet* dilakukan dalam kurun waktu Desember 2022 hingga Januari 2023. Data diambil pada Januari 2023 dikarenakan adanya regulasi terbaru dari X Corp. yang merubah limit pengambilan data menjadi 1.500 *tweet* perbulan.
3. Analisis sentimen hanya akan dilakukan pada *tweet* yang ditulis dalam bahasa Indonesia.
4. Metode yang digunakan dalam penelitian ini adalah algoritma *Extreme Gradient Boosting*.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan penelitian ini sebagai berikut:

1. Mengembangkan model klasifikasi sentimen menggunakan algoritma XGBoost yang optimal untuk memprediksi sentimen *tweet* dengan kata kunci "indihome" secara akurat.
2. Menganalisis hasil klasifikasi sentimen untuk mengidentifikasi pola dan karakteristik sentimen yang dominan terkait layanan IndiHome.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah:

1. Mengetahui tingkat akurasi dan mengembangkan metode klasifikasi yang optimal dengan menggunakan algoritma eXtreme Gradient Boosting dalam mengklasifikasi sentimen dari layanan Indihome.
2. Mengembangkan *hyperparameter tuning* yang optimal dengan konfigurasi *learning rate*, *max depth*, dan *n_estimator* terbaik.
3. Menganalisa dan mengklasifikasikan sentimen dari layanan Indihome yang berupa pendapat pengguna pada sosial media Twitter(X).

1.6 Metodologi Penelitian

Tahapan dalam proses penelitian ini adalah:

1. *Data Collection & Data Cleaning*

Pada tahap ini dilakukan pengambilan data, kemudian data dibersihkan dari atribut yang tidak akan digunakan.

2. *Data Preprocessing*

Tahapan ini meliputi *text pre-processing*, *labelling sentiment*, dan ekstraksi fitur menggunakan TF-IDF.

3. *Modelling & Evaluation*

Pada tahapan ini dilakukan pengaplikasian model XGBoost dan juga konfigurasi parameter lainnya hingga menghasilkan sebuah informasi berupa sentimen dan kinerja model.

4. *Analysis*

Analisis dilakukan untuk mengetahui kinerja model dan kelas sentimen yang paling sering muncul.

1.7 Sinematika Penulisan

Sistematika dalam penulisan ini bertujuan untuk mempermudah pembacaan dan gambaran umum tentang penelitian yang akan dilakukan dengan sistematika penulisan penelitian sebagai berikut:

BAB I PENDAHULUAN

Bab ini akan membahas tentang latar belakang masalah, rumusan, masalah, batasan penelitian, tujuan penelitian, manfaat penelitian, metode penelitian, dan sistematika penelitian.

BAB II LANDASAN TEORI

Bab ini akan membahas mengenai berbagai teori yang mendasari dari analisis permasalahan yang berhubungan dengan topik yang dibahas.

BAB III METODOLOGI

Bab ini akan membahas mengenai metodologi yang digunakan dalam penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Bab ini akan membahas hasil dari proses dalam penelitian, mulai dari pengolahan data hingga hasil dari penerapan model algoritma.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi tentang kesimpulan yang merupakan rangkuman dari hasil pembahasan masalah pada penelitian dan saran yang perlu diperhatikan berdasarkan kekurangan yang ada dalam penelitian.