



Analisis Sentiment Review Produk Tokopedia Menggunakan Ensemble Learning

Fatwa Sifaun Nahar^{*1}, Fetty Tri Anggraeny², Henni Endah Wahanani³

^{1,2,3}Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur, Surabaya, Indonesia

Email: fatwanahar@gmail.com^{*1}; fettyanggraeny.if@upnjatim.ac.id²; henniendah@upnjatim.ac.id³

Nahar, F. S., Anggraeny, F. T., & Wahanani, H. E. (2026). Analisis Sentiment Review Produk Tokopedia Menggunakan Ensemble Learning. *Journal Cerita: Creative Education of Research in Information Technology and Artificial Informatics*, 12(1), 12-21

DOI: <https://doi.org/10.33050/3xhrfn68>

ABSTRAK

Perkembangan pesat e-commerce di Indonesia mendorong adanya analisis ulasan produk untuk meningkatkan kualitas layanan dan memfasilitasi pengambilan keputusan konsumen. Salah satu platform e-commerce yang populer, Tokopedia, mengandalkan ulasan produk sebagai sarana untuk membangun kepercayaan dan mempengaruhi keputusan pembelian. Namun, dengan banyaknya ulasan yang tersedia, analisis manual menjadi tidak efisien dan rawan kesalahan. Oleh karena itu, penelitian ini bertujuan untuk menerapkan analisis sentimen pada ulasan produk di Tokopedia dengan menggunakan algoritma pembelajaran mesin berbasis ensemble learning, yaitu Decision Tree, Random Forest, AdaBoost, dan XGBoost. Penelitian ini dimulai dengan pengambilan dataset ulasan produk dari toko Rexus di Tokopedia, yang terdiri dari 470 ulasan yang kemudian diproses melalui tahap pra-pemrosesan seperti pembersihan teks, penghapusan stopword, stemming, pelabelan dengan VADER, dan ekstraksi fitur menggunakan TF-IDF. Secara akurasi yang terbaik dari keempat model tersebut adalah Random Forest dan AdaBoost dengan akurasi sebesar 81%, lalu XGBoost dengan akurasi sebesar 80%, dan terakhir yaitu Decision Tree dengan akurasi yaitu 79%.

Kata kunci: Analisis Sentimen, Decision Tree, Random Forest, AdaBoost, XGBoost

ABSTRACT

The rapid development of e-commerce in Indonesia has encouraged product review analysis to improve service quality and facilitate consumer decision-making. One of the popular e-commerce platforms, Tokopedia, relies on product reviews as a means to build trust and influence purchasing decisions. However, with the large number of reviews available, manual analysis becomes inefficient and error-prone. Therefore, this study aims to apply sentiment analysis to product reviews on Tokopedia using ensemble learning-based machine learning algorithms, namely Decision Tree, Random Forest, AdaBoost, and XGBoost. This study began by taking a product review dataset from the Rexus store on Tokopedia, consisting of 470 reviews which were then processed through pre-processing stages such as text cleaning, stopword removal, stemming, labeling with VADER, and feature extraction using TF-IDF. In terms of accuracy, the best of the four models is Random Forest and AdaBoost with an accuracy of 81%, then XGBoost with an accuracy of 80%, and finally Decision Tree with an accuracy of 79%.

Keywords: *Sentiment Analysis, Decision Tree, Random Forest, AdaBoost, XGBoost.*

I. PENDAHULUAN

Pasar daring di Indonesia telah mengalami peningkatan yang signifikan, seiring dengan pesatnya perkembangan teknologi (Muhamad Arif Saifudin, 2024). Hal ini menjelaskan mengapa sebagian besar masyarakat Indonesia lebih memilih berbelanja secara online daripada berbelanja langsung di toko fisik. Menurut data dari Katadata, penggunaan e-commerce di Indonesia mencapai 88,1 persen, menempatkan negara ini di posisi teratas dalam daftar negara dengan persentase penggunaan e-commerce tertinggi di dunia (Lidwina, 2021). Pertumbuhan pasar online ini juga mendorong banyak penjual untuk mengalihkan bisnis mereka ke platform digital. Tokopedia, misalnya, merupakan salah satu aplikasi e-commerce yang populer di Indonesia. Berdasarkan data dari Katadata, Tokopedia menjadi platform online terpopuler kedua setelah Shopee, dengan pengguna mencapai 117,03 juta (Ahdiat, 2023).

Aplikasi Tokopedia menawarkan berbagai produk yang bisa dibeli secara online (Aulia & Hermawan, 2023). Meskipun platform e-commerce ini memberikan banyak kemudahan, ada beberapa kekurangan, salah satunya adalah minimnya interaksi langsung dengan produk (Siringoringo & Jamaluddin, 2019). Untuk mengatasi hal ini, fitur ulasan produk diperkenalkan. Ulasan menjadi alat komunikasi yang efektif dalam pemasaran dan iklan, serta berperan penting dalam proses pengambilan keputusan konsumen untuk menilai dan memilih produk yang akan dibeli. Ulasan produk memungkinkan konsumen untuk mengetahui kualitas suatu barang berdasarkan pengalaman konsumen lain. Selain itu, ulasan ini juga penting dalam membangun kepercayaan konsumen

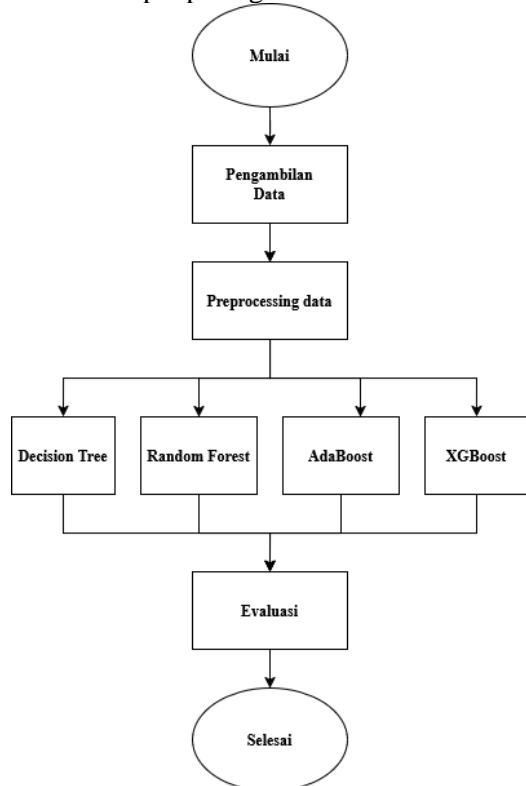
terhadap produk yang ingin mereka beli (Sera, Hazriani, Mirfan, & Yuyun, 2023).

Dengan meningkatnya jumlah ulasan produk di Tokopedia, tugas untuk mengatur dan memilah ulasan menjadi lebih kompleks dan memakan waktu jika dilakukan secara manual. Situasi ini bisa mempengaruhi kualitas produk karena penjual mungkin tidak menyadari mana produk yang mendapatkan ulasan positif atau negatif (Saputra, Adiwijaya, & Mubarak, 2017). Oleh karena itu, diperlukan metode otomatis untuk menganalisis ulasan tersebut, yang dikenal sebagai analisis sentimen. Analisis sentimen adalah metode yang digunakan untuk mendeteksi, mengekstraksi, dan mengolah data teks guna mengungkap emosi yang terkandung dalam suatu pernyataan atau pendapat (Safira & Hasan, 2023).

Dalam konteks analisis sentimen, penggunaan algoritma pembelajaran mesin ensemble learning seperti Extreme Gradient Boosting (XGBoost), Decision Tree, Random Forest, dan juga AdaBoost telah terbukti efektif dalam menangani data yang berdimensi tinggi dan kompleks (Elzbieta Turska, 2021). Untuk pengolahan bahasa alami (NLP), teknik Term Frequency-Inverse Document Frequency (TF-IDF) sering digunakan untuk mengekstrak fitur penting dari teks dengan mempertimbangkan frekuensi kata relatif terhadap keunikannya di seluruh dokumen (Ramos, 2003). Selain itu, Valence Aware Dictionary and sEntiment Reasoner (VADER) adalah alat yang dirancang khusus untuk analisis sentimen, terutama untuk teks dari media sosial, dengan menggunakan leksikon yang dikurasi untuk mengklasifikasikan teks ke dalam kategori sentimen positif, negatif, atau netral (Hutto & Gilbert, 2014).

II. METODE PENELITIAN

Untuk metode penelitian ada beberapa tahapan yang akan dilakukan, tahapan tahapan tersebut terdapat pada gambar nomor 1



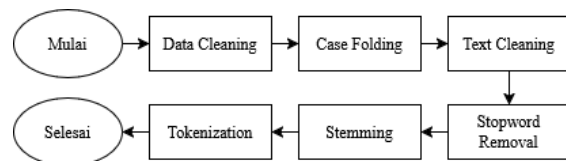
Gambar 1. Tahapan Penelitian

A. Pengambilan Data

Pengambilan data akan dilakukan di website Tokopedia. Toko yang akan diambil reviewsnya adalah toko rebus dimana toko tersebut menyediakan beberapa produk gaming seperti, keyboard gaming, mouse gaming, headset gaming, dll. Untuk link dari toko adalah berikut <https://www.tokopedia.com/rebusid/review/>. Pengambilan data akan digunakan library python yaitu selenium. Library selenium akan mengambil data reviews dari produk produk yang ada di rebus official store. Untuk jumlah data yang diambil berjumlah 470 data dikarenakan Tokopedia membatasi jumlah data yang boleh diambil yaitu maksimal 50 halaman dimana setiap halaman berisikan 10 ulasan. Data yang diperoleh masih kotor karena banyak mengandung emotikon, tanda baca, dll. Selain itu juga dataset yang ada masih belum ada pelabelan jadi harus diberikan label terlebih dahulu. Untuk pelabelan akan digunakan library Valence Aware Dictionary and sEntiment Reasoner (VADER).

B. Preprocessing Data

Untuk Preprocessing sendiri ada beberapa tahap yang akan dilakukan yaitu menghilangkan data kosong dan data duplikat, menghapus karakter non huruf ataupun angka, case folding yaitu merubah semua huruf menjadi kecil, stopword removal, pelabelan menggunakan VADER. stemming atau merubah semua kata menjadi bentuk dasarnya, tokenization, dan yang terakhir adalah TF-IDF. Tahapan bisa dilihat di gambar 2.



Gambar 2. Preprocessing Data

1. Data Cleaning

Tahapan pertama pada preprocessing yaitu menghilangkan atau menghapus data data yang kosong ataupun data yang terduplikat. Tujuan dari pembersihan ini yaitu agar data yang kosong tidak menjadi beban saat dilakukan perhitungan mesin. Selain itu, jika ada yang terduplikasi maka akan menyebabkan terjadinya noise pada dataset dimana perhitungan mesin yang dilakukan bisa saja terjadi bias karena ada data yang seharusnya hanya muncul satu kali namun didalam dataset muncul dua atau bahkan lebih.

2. Case Folding

Karena data diambil dari media sosial, gaya penulisan teks pasti akan sangat bervariasi. Case Folding adalah langkah dalam pemrosesan teks yang melibatkan konversi semua karakter dalam teks menjadi huruf kecil. Hanya karakter yang berada dalam rentang huruf 'a' hingga 'z' yang akan mengalami perubahan ini

3. Pembersihan Teks

Data yang diambil dari web browser sering kali tidak bersih dan mengandung banyak karakter non-alfabet. Karakter non-alfabet seperti "?", "!", atau karakter lainnya seperti emotikon dan teks HTML dapat mempengaruhi pembobotan, sehingga diperlukan proses text cleaning. Text cleaning adalah proses pembersihan dokumen dan penyaringan kata-kata yang tidak diperlukan, seperti HTML, emotikon, hashtag, mention, dan URL

4. Stopword Removal

Stopword removal adalah proses menghapus kata-kata yang tidak penting yang dapat mempengaruhi proses klasifikasi. Contoh stopwords dalam bahasa Indonesia meliputi “di”, “ke”, “dari”. Menghilangkan stopwords atau kata-kata yang tidak penting ini bermanfaat untuk mempercepat proses eksekusi data pada sistem. Terdapat banyak kamus stopwords yang tersedia di internet, salah satunya adalah kamus stopwords dari Natural Language Toolkit (NLTK) yang akan digunakan dalam penelitian ini.

5. Stemming

Stemming merupakan bagian dari proses prapemrosesan data, di mana teknik ini digunakan untuk mengubah kata-kata ke bentuk dasarnya atau ke bentuk kata kerja infinitif. Sebagai contoh, kata-kata seperti "menyanyikan," "menyanyi," dan "nyanyi" dapat disederhanakan menjadi bentuk dasar "nyanyi." Proses ini membantu mengurangi variasi kata yang memiliki makna serupa. Algoritma yang akan digunakan dalam penelitian ini adalah Sastrawi, yaitu sebuah kamus atau perpustakaan yang digunakan untuk melakukan stemming pada kata-kata dalam bahasa Indonesia.

6. Tokenization

Tokenisasi adalah proses memecah urutan karakter menjadi bagian-bagian yang disebut token, yang bisa berupa kata atau frasa. Tujuan dari proses tokenisasi ini adalah untuk membagi teks menjadi unit-unit yang lebih mudah diolah, seperti kata atau frasa.

C. Pelabelan Data

Pelabelan data akan dilakukan secara otomatis menggunakan library Valence Aware Dictionary and Sentiment Reasoner (VADER). Proses pelabelan ini menghasilkan tiga kategori: netral, positif, dan negatif. Kategori positif menunjukkan bahwa pelanggan merasa puas dan senang dengan produk yang dibeli. Kategori netral mengindikasikan bahwa pelanggan tidak merasa puas atau tidak puas dengan produk yang diterima. Sementara itu, kategori negatif menunjukkan ketidakpuasan terhadap pelayanan atau produk yang dibeli.

VADER bekerja dengan memberikan Compound Score pada sebuah kalimat. Compound Score ini dihitung dari rata-rata penilaian setiap kata dalam teks, berdasarkan apakah kata tersebut bernilai negatif atau positif, dengan masing-masing memiliki bobot penilaian

tertentu. Setelah Compound Score diperoleh, klasifikasi dilakukan berdasarkan nilai-nilai tersebut: jika Compound Score lebih dari 0.5, teks diklasifikasikan sebagai positif; jika kurang dari -0.5, diklasifikasikan sebagai negatif; dan jika nilainya berada dalam rentang antara -0.5 hingga 0.5 (termasuk), teks akan diklasifikasikan sebagai netral.

D. Ekstraksi Fitur

Pembobotan term atau term weighting adalah langkah penting dalam meningkatkan kemampuan analisis sentimen dalam text mining. Dalam penelitian ini, digunakan pendekatan Term Frequency-Inverse Document Frequency (TF-IDF). Term Frequency ($tf(w,d)$) dihitung berdasarkan frekuensi kemunculan suatu kata dalam sebuah teks atau dokumen. Semakin sering kata tersebut muncul, semakin tinggi bobot yang diberikan. Di sisi lain, Inverse Document Frequency (IDF) berfungsi sebagai metode pembobotan untuk mengukur sejauh mana kemunculan suatu token dalam keseluruhan kumpulan teks [18]. Dengan menggabungkan TF dan IDF, dapat diperoleh bobot yang lebih akurat dan informatif untuk setiap kata dalam teks, sehingga meningkatkan kualitas analisis sentimen pada proses text mining. Langkah-langkah dalam TF-IDF antara lain [19]

1. Menghitung Term Frequency (TF)

TF dihitung untuk setiap kata dalam sebuah dokumen. Ini menunjukkan frekuensi kemunculan kata dalam dokumen relatif terhadap jumlah total kata. Rumus TF sendiri adalah:

$$TF(t_k, d_j) = f(t_k, d_j) \quad (1)$$

di mana $f(t_k, d_j)$ mendefinisikan jumlah kemunculan term k dalam dokumen j .

2. Menghitung Document Frequency (DF)

DF adalah jumlah dokumen dalam korpus yang mengandung kata tersebut. DF memberikan indikasi seberapa umum atau jarang kata tersebut muncul di seluruh dokumen.

3. Menghitung Inverse Document Frequency (IDF)

Langkah berikutnya adalah menghitung seberapa informatif suatu kata dengan mengukur seberapa sering kata tersebut muncul dalam keseluruhan koleksi dokumen. Rumusnya adalah

$$IDF(t_k) = \log \frac{D}{df(t)} \quad (2)$$

mana D adalah jumlah dokumen dalam dataset, dan $df(t)$ adalah jumlah dokumen yang mengandung term tersebut.

4. Menghitung Skor TF-IDF

Terakhir, skor TF-IDF untuk setiap kata dalam dokumen dihitung dengan mengalikan TF dan IDF. Ini dilakukan untuk memberikan bobot yang menyesuaikan frekuensi kata dalam dokumen dengan seberapa umum atau jarang kata tersebut muncul dalam korpus. Rumus TF-IDF adalah:

$$TF\ IDF(t_k, d_j) = TF(t_k, d_j) * IDF(t_k) \quad (3)$$

E. Klasifikasi

Untuk klasifikasi akan dilakukan dengan beberapa pemodelan.

1. Decision Tree

Pohon keputusan adalah algoritma pembelajaran mesin yang populer yang digunakan untuk tugas klasifikasi dan regresi. Algoritma ini bekerja dengan membagi dataset secara rekursif menjadi cabang-cabang berdasarkan fitur tertentu, yang kemudian membentuk model seperti pohon dari keputusan-keputusan. Setiap node internal dalam pohon mewakili suatu tes pada fitur tertentu, setiap cabang mewakili hasil dari tes tersebut, dan setiap node daun mewakili label kelas atau hasil akhir

2. Random Forest

Random Forest adalah algoritma pembelajaran mesin yang menggabungkan banyak pohon keputusan untuk meningkatkan akurasi prediksi dan mengurangi varians. Algoritma ini bekerja dengan membangun beberapa pohon keputusan pada subset acak dari dataset dan kemudian menggabungkan hasilnya untuk membuat prediksi yang lebih stabil dan akurat. Random Forest dikenal efektif dalam menangani data yang kompleks dan tinggi dimensi

3. AdaBoost

AdaBoost adalah algoritma ensemble yang digunakan dalam pembelajaran mesin untuk meningkatkan performa model klasifikasi. Algoritma ini bekerja dengan menggabungkan beberapa model lemah (weak learners), biasanya pohon keputusan sederhana, dan menekankan pada data yang sulit diklasifikasikan oleh model sebelumnya. Setiap model berturut-turut fokus pada kesalahan yang dibuat oleh model sebelumnya, sehingga menghasilkan model akhir yang lebih akurat

4. XGBoost

XGBoost adalah algoritma yang sangat efektif dalam pembelajaran mesin, dikenal karena kemampuannya mengatasi masalah overfitting dan performa tinggi dalam tugas-tugas klasifikasi dan regresi. Algoritma ini menggunakan teknik boosting yang secara bertahap memperbaiki kesalahan model sebelumnya, sehingga meningkatkan akurasi prediksi. XGBoost juga sangat efisien dalam penggunaan sumber daya komputasi dan mampu menangani data yang tidak seimbang dengan baik

F. Evaluasi

Untuk Evaluasi akan dilakukan pengujian menggunakan confusion matrix untuk mengetahui akurasi dan f1-score dari masing masing label class. Selain itu, akan dilakukan juga dengan cara membandingkan f1-score dan akurasi sehingga bisa mengetahui pemodelan mana yang lebih baik.

III. HASIL DAN PEMBAHASAN

Pada penelitian sentiment analysis menggunakan ensemble learning ini, terdapat empat skenario pengujian model yaitu pengujian menggunakan algoritma decision tree, random forest, AdaBoost dan XGBoost. Skenario ini untuk mengetahui peforma paling baik untuk sentiment analysis review produk tokopedia, Berikut merupakan hasil dari skenario yang dilakukan

A. Pengumpulan Data

Untuk contoh hasil data yang telah diperoleh bisa dilihat Digambar 3

Ulasan			
ok, mantab			
Kualitas bagus, berfungsi dengan normal.			
packing ancurrr			
Barang sudah sampai dan sesuai, terima kasih			
Barang sudah sampai dan sesuai, terima kasih			
Barang sudah sampai dan sesuai, terima kasih			
Mantap			
mantap banget jir dingin banget sumpah kek di ai			
barang sesuai dgn deskripsi, keyboard yang cocok			
Mantap berfungsi dengan baik, kualitas Origina p			
mantap barang bekerja dgn baik, semoga awet			
sesuai ekspektasi 5Yr— keren paraaaaaaahhhh			

Gambar 3. Data Mentah

Data yang diambil berjumlah 470 data dan jika dilihat dari gambar 3, data masih sangat tidak

layak untuk masuk kedalam machine learning, jadi harus dibersihkan terlebih dahulu dengan preprocessing data agar siap untuk dimasukkan kedalam machine learning.

B. Preprocessing Data

Sebelum masuk kedalam pembelajaran mesin terlebih dahulu data akan dibersihkan dari beberapa hal yakni data terduplikat, data kosong, ketidakstabilan huruf besar maupun huruf kecil, tanda baca, emotikon, angka. Selanjutnya data akan dilakukan stopwords removal dan stemming. Lalu, dilakukan pelabelan dengan VADER dan yang terakhir ada tokenization dan ekstraksi fitur menggunakan TF-IDF.

1. Data Cleaning

Pertama data akan dilakukan cleaning, dimana data duplikat dan data kosong akan dihilangkan. Tahap ini penting karena jika ada data duplikat maka akan membuat machine learning bias ke data yang duplikat tersebut. Selain itu, dilakukan penghapusan data kosong atau NaN karena data kosong tidak memberikan nilai kedalam machine learning sehingga lebih baik dihilangkan saja.

2. Case Folding

Tahap selanjutnya yaitu Case folding dimana huruf akan diubah menjadi huruf kecil semua.

Tabel 1. Hasil Case Folding

Data Sebelum	Data Sesudah
Mantap berfungsi dengan baik, kualitas Originala pula	mantap berfungsi dengan baik, kualitas origina pula
â • â • â • â • â •	â • â • â • â • â •

Pada Tabel 1 menampilkan hasil sebelum dan sesudah data menjalani proses Case Folding di mana semua karakter akan diubah menjadi huruf kecil guna mempermudah konsistensi dalam pembelajaran mesin. Dapat dilihat pada tabel bahwa di awal kalimat terdapat huruf besar, tetapi setelah melewati langkah ini, semua teks akan diubah menjadi huruf kecil.

3. Text Cleaning

Selanjutnya yaitu tahapan text cleaning dimana karakter seperti emotikon akan dihapus.

Tabel 2. Hasil Text Cleaning

Data Sebelum	Data Sesudah
mantap berfungsi dengan baik, kualitas origina pula ★★★★★	mantap berfungsi dengan baik kualitas original pula

Dalam Tabel 2, ditampilkan hasil dari langkah pembersihan, di mana teks akan dihilangkan dari berbagai jenis tanda baca, emotikon, dan angka. Tujuannya adalah untuk mengurangi gangguan dalam dataset supaya machine learning dapat bekerja dengan lebih tepat. Tabel 2 menunjukkan adanya beberapa tanda baca serta emotikon bintang, namun setelah melewati proses ini, teks tampak bebas dari tanda baca maupun emotikon.

4. Stopword removal

Tahapan selanjutnya yaitu stopwords removal dimana kata kata penghubung seperti di ke dari akan dihapus.

Tabel 3. Hasil stopwords removal

Data Sebelum	Data Sesudah
mantap berfungsi dengan baik kualitas original pula	mantap berfungsi kualitas original

Dalam Tabel 3 menampilkan hasil data setelah melalui stopwords removal. sebagai contoh, pernyataan "mantap berfungsi dengan baik kualitas original pula" setelah melalui proses Penghapusan Stopword akan menjadi "mantap berfungsi kualitas original" yang lebih singkat tetapi tetap menjaga inti maknanya.

5. Stemming

Tahap selanjutnya yaitu stemming dimana setiap kata akan diubah menjadi kata dasarnya.

Tabel 4. Hasil Stemming

Data Sebelum	Data Sesudah
mantap berfungsi kualitas original	mantap fungsi kualitas original

Pada tabel 4 terlihat kalimat yang sudah melalui stemming. Proses ini bertujuan untuk menyederhanakan kata-kata menjadi bentuk dasar mereka, sehingga memudahkan analisis teks dengan mengurangi variasi kata yang berasal dari bentuk yang sama. Sebagai

contoh, kata "berfungsi" diubah menjadi "fungsi," sehingga kalimat "kualitas bagus berfungsi original" menjadi "kualitas bagus fungsi original."

6. Tokenization

Selanjutnya yaitu tahap tokenization dimana sebuah kalimat akan dipecah menjadi beberapa kata.

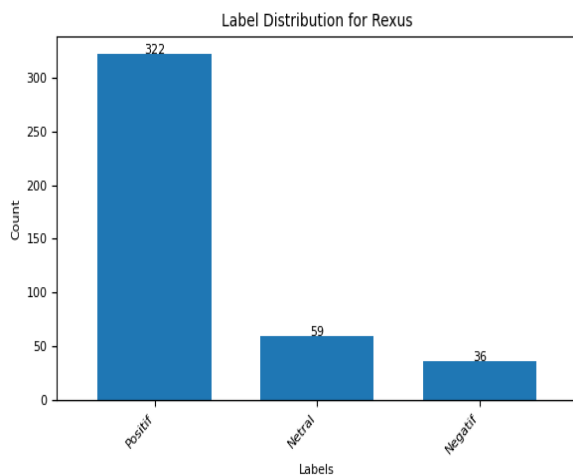
Tabel 5. Hasil Tokenization

Data Sebelum	Data Sesudah
mantap fungsi kualitas original	['mantap', 'fungsi', 'kualitas', 'original']

Pada tabel 5 terlihat contoh hasil tokenization dimana kalimat yang awalnya utuh akan dipecah menjadi kata kata guna memudahkan machine learning untuk memproses data.

C. Pelabelan Data

Setelah melalui data Preprocessing seperti case folding, pembersihan teks, stopword removal, stemming, pelabelan, dan tokenization, selanjutnya akan dilakukan pelabelan data. Pelabelan data yaitu memberikan class kepada masing masing kalimat dalam dataset dan akan dikategorikan menjadi 3 class yaitu positif, netral, dan negatif. Library pelabelan yang digunakan yaitu Valence Aware Dictionary and Sentiment Reasoner (VADER). Untuk hasil pelabelan bisa dilihat di gambar 4.



Gambar 4. Distribusi Label

Dari data tersebut dapat dilihat persebaran label class data yang ada. Untuk review positif berjumlah 322, untuk review netral berjumlah 59, dan untuk review negatif berjumlah 36. Dari Chart tersebut menunjukkan data berjumlah 417 yang dimana di

awal data mentah berjumlah 470. Hal ini dikarenakan beberapa data ada yang terduplikasi ataupun ada data yang kosong sehingga data terpaksa harus dihapus agar tidak memengaruhi pembelajaran mesin.

D. Ekstraksi Fitur

Setelah melakukan pelabelan selantunya yaitu dilakukan ekstraksi fitur. fitur diambil dengan memanfaatkan TF-IDF untuk menilai seberapa penting suatu kata di dalam dokumen dibandingkan dengan seluruh dataset. Melalui proses stemming, kata-kata diringkas menjadi bentuk aslinya, sehingga dianggap sebagai satu fitur. TF-IDF memberikan nilai yang lebih besar pada kata-kata yang muncul sering dalam satu dokumen tetapi jarang di dokumen lainnya, yang membantu model berkonsentrasi pada informasi yang relevan dan meningkatkan ketepatan analisis teks.

abisssss	action	ad	admin	adminnya	aga	aint	air	aja	akalin	ake
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.2412649765083988	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0

Gambar 5. Contoh Ekstraksi Fitur

Gambar 5 memperlihatkan hasil dari pengambilan fitur menggunakan pendekatan TF-IDF. Nilai-nilai yang terdapat dalam tabel ini merupakan angka TF-IDF yang mengindikasikan pentingnya sebuah kata dalam konteks dokumen tertentu, jika dibandingkan dengan keseluruhan kumpulan dokumen. Sebagai contoh, kata "admin" yang terdapat pada baris keenam memiliki nilai TF-IDF 0.241, yang menunjukkan bahwa kata tersebut cukup berpengaruh dalam dokumen tersebut, namun tidak terlalu sering muncul di dokumen lainnya.

E. Klasifikasi dan Evaluasi

Selanjutnya yaitu klasifikasi dan evaluasi, namun sebelum itu pembagian data akan dilakukan dimana data akan dibagi menjadi data training dan data testing. Untuk data akan akan dibagi menjadi 80:20 dimana data training berjumlah 80% dan data test akan berjumlah

20%. Untuk 4 uji coba dari masing masing model akan dijelaskan dibawah.

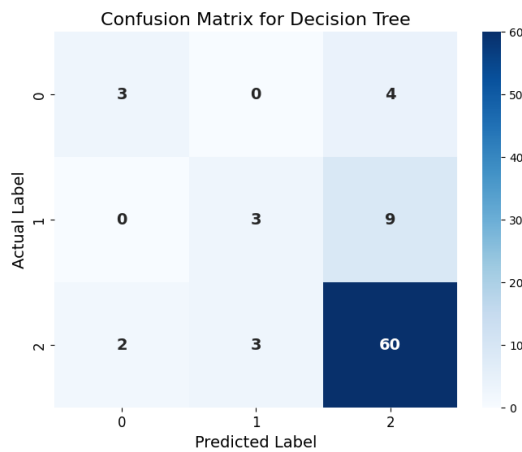
1. Decision Tree

Pemodelan decision tree dilakukan dengan hyperparameter default dan untuk hasilnya bisa dilihat di tabel 6

Tabel 6. Classification Report Model DT

Parameter Uji Coba	Hasil
Akurasi	0.79%
Weighted Avg Precision	0.76%
Weighted Avg Recall	0.79%
Weighted Avg F1-Score	0.76%

Untuk Confusion Matrix yang dihasilkan yaitu



Gambar 5. Confusion Matrix DT

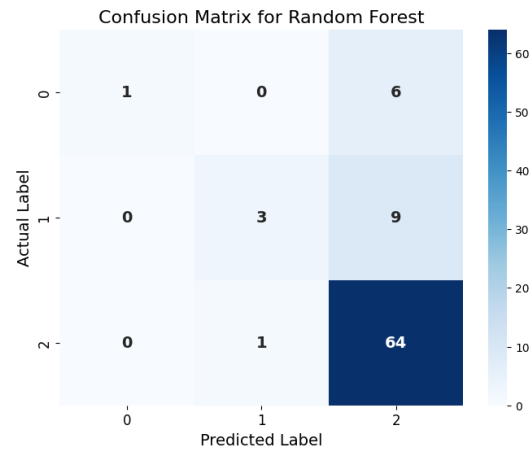
2. Random Forest

Pemodelan random forest dilakukan dengan hyperparameter default dan untuk hasilnya bisa dilihat di tabel 7

Tabel 7. Classification Report Model RF

Parameter Uji Coba	Hasil
Akurasi	0.81%
Weighted Avg Precision	0.82%
Weighted Avg Recall	0.81%
Weighted Avg F1-Score	0.76%

Untuk Confusion Matrix yang dihasilkan yaitu



Gambar 6. Confusion Matrix RF

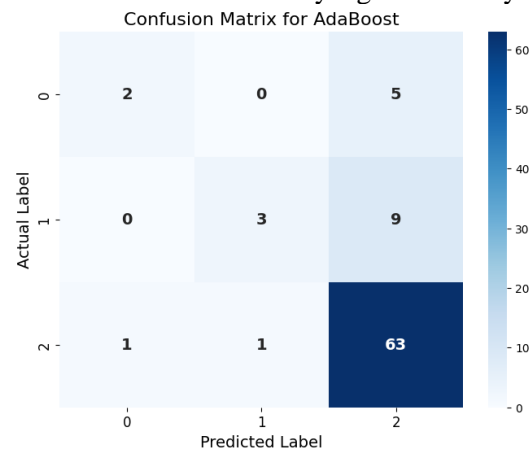
3. AdaBoost

Pemodelan AdaBoost dilakukan dengan hyperparameter default dan untuk hasilnya bisa dilihat di tabel 8

Tabel 8. Classification Report Model AdaBoost

Parameter Uji Coba	Hasil
Akurasi	0.81%
Weighted Avg Precision	0.80%
Weighted Avg Recall	0.81%
Weighted Avg F1-Score	0.77%

Untuk Confusion Matrix yang dihasilkan yaitu



Gambar 5. Confusion Matrix AdaBoost

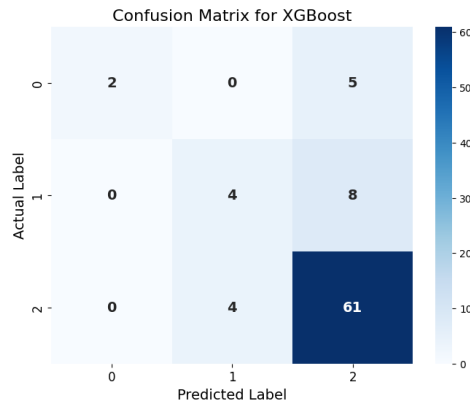
4. XGBoost

Pemodelan random forest dilakukan dengan hyperparameter default dan untuk hasilnya bisa dilihat di tabel 9

Tabel 9. Classification Report Model XGBoost

Parameter Uji Coba	Hasil
Akurasi	0.80%
Weighted Avg Precision	0.79%
Weighted Avg Recall	0.80%
Weighted Avg F1-Score	0.77%

Untuk Confusion Matrix yang dihasilkan yaitu



Gambar 7. Confusion Matrix XGBoost

IV. KESIMPULAN

Berdasarkan hasil perbandingan kinerja model yang telah dihitung, berikut adalah beberapa poin kesimpulan yang dapat diambil:

1. Random Forest unggul dalam akurasi yang mencapai 81%, tetapi cenderung memiliki masalah dalam mendeteksi kelas 0 dan 1 dengan benar, meskipun cukup kuat dalam mendeteksi kelas mayoritas (kelas 2).
2. AdaBoost dan XGBoost memiliki performa yang cukup baik dengan akurasi AdaBoost mencapai 81% dan XGBoost mencapai 80%, dengan AdaBoost sedikit lebih baik dalam keseimbangan recall dan precision, tetapi keduanya juga memiliki tantangan serupa dalam mengklasifikasikan kelas 0 dan 1.
3. Decision Tree memiliki keakuratan yang sedikit lebih rendah, dengan masalah yang lebih besar dalam mendeteksi kelas 0 dan 1, meskipun masih cukup baik dalam mendeteksi kelas mayoritas.
4. Secara keseluruhan, meskipun ada perbedaan kinerja di antara model-model ini, semua model menunjukkan performa yang sangat baik dalam memprediksi kelas mayoritas (kelas 2), tetapi masing-masing memiliki

kesulitan dalam mengenali kelas minoritas (kelas 0 dan 1).

DAFTAR PUSTAKA

- Ahdiat, A. (2023, Mei 3). databoks. (Kata Data Media Network) Dipetik Maret 7, 2024, dari <https://databoks.katadata.co.id/datapublish/2023/05/03/5-e-commerce-dengan-pengunjung-terbanyak-kuartal-i-2023>
- Aulia, M., & Hermawan, A. (2023). Analisis Perbandingan Algoritma SVM, Naïve Bayes, dan Perceptron untuk Analisis Sentimen Ulasan Produk Tokopedia. *JURNAL MEDIA INFORMATIKA BUDIDARMA*.
- Elzbieta Turska, S. J. (2021). Mood Disorder Detection in Adolescents by Classification Trees, Random Forests and XGBoost in Presence of Missing Data. *Entropy*. doi:<https://doi.org/10.3390/e23091210>
- Gifari, O. I., Adha, M., Hendrawan, I. R., & Durrand, F. F. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine. *JIFOTECH (JOURNAL OF INFORMATION TECHNOLOGY)* Vol. 2, No. 1, Maret 2022.
- Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*.
- Lidwina, A. (2021, Juni 4). Databoks. Dipetik Maret 7, 2024, dari Katadata Media Network: <https://databoks.katadata.co.id/datapublish/2021/06/04/penggunaan-e-commerce-indonesia-tertinggi-di-dunia>
- Muhamad Arif Saifudin, H. E. (2024). IMPLEMENTASI ALGORITMA ASOSIASI FP-GROWTH DAN KLASIFIKASI K-MEANS TERHADAP POLA PEMBELIAN KONSUMEN DI MARKETPLACE SHOPEE. *Jurnal Mahasiswa Teknik Informatika*, 764-771.
- Nabila Ayu Puspita, F. T. (2024). ANALISIS SENTIMEN DENGAN ALGORITMA NAÏVE BAYES CLASSIFIER TERHADAP ULASAN APLIKASI MY

- F&B ID. Jurnal Mahasiswa Teknik Informatika, 1090-1095.
- Ramos, J. (2003). Using TF-IDF to Determine Word Relevance in Document Queries. In Proceedings of the First Instructional Conference on Machine Learning.
- Safira, A., & Hasan, F. N. (2023). ANALISIS SENTIMEN MASYARAKAT TERHADAP PAYLATER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER. *Zonasi Jurnal Informasi*.
- Saputra, A. D., Adiwijaya, & Mubarak, M. S. (2017). Klasifikasi Sentimen Pada Level Aspek Terhadap Ulasan Produk Berbahasa Inggris Menggunakan Bayesian Network (case Study : Data Ulasan Produk Amazon). *e-Proceeding of Engineering : Vol.4, No.3 Desember 2017*.
- Sera, E., Hazriani, Mirfan, & Yuyun. (2023). Analisis Sentimen Ulasan Produk di E-Commerce Bukalapak Menggunakan Natural Language Processing. Prosiding Seminar Nasional Sistem Informasi dan Teknologi (SISFOTEK) ke 7 Tahun 2023.
- Siringoringo, R., & Jamaluddin. (2019). Text Mining dan Klasterisasi Sentimen Pada Ulasan Produk Toko Online. *J. Teknol. dan Ilmu Komput. Prima*, vol. 2, no. 1., 2019, 41–48.