

Classification of Consumer Sentiment towards Muslim Fashion Products Using TF-IDF and Machine Learning

Rahmayani¹⁾, Irfan Nurudin^{2*)}

¹⁾²⁾Informatika, STMIK El Rahma Yogyakarta

^{*)}Correspondence author: irfannurudin1@gmail.com, Yogyakarta, Indonesia

DOI: <https://doi.org/10.37012/jtik.v12i1.3777>

Abstract

The growth of e-commerce has driven an increase in transactions for Muslim fashion products, generating a vast number of consumer reviews containing information about product experiences and perceptions. Analyzing these reviews is crucial, as sentiment information allows for a more systematic understanding of consumer satisfaction trends compared to manual analysis. This study aims to classify consumer sentiment regarding Muslim fashion products and compare the performance of several machine learning algorithms in this task. The research dataset was obtained from PRDECT-ID, initially comprising 5,400 reviews; filtering for the Muslim fashion category yielded 200 reviews. Sentiments were determined based on review content, and "Neutral" labeled data were excluded, resulting in a final set of 135 reviews (97 negative and 38 positive). The research process involved text preprocessing, feature weighting using Term Frequency-Inverse Document Frequency (TF-IDF), classification using Multinomial Naive Bayes (MNB), Logistic Regression (LR), and Support Vector Machine (SVM), and evaluation using Stratified 5-Fold Cross-Validation. This study contributes an empirical comparison of the three algorithms applied to Muslim fashion product reviews, accounting for class imbalance and multiple evaluation metrics. Experimental results indicate that SVM achieved the best performance, with an accuracy of 84.44%, precision of 81.43%, recall of 60.36%, and an F1-score of 68.00%. These findings demonstrate that the combination of TF-IDF and SVM is more effective than MNB and LR for sentiment classification on the study's dataset.

Keywords: Muslim Fashion, Sentiment Classification, TF-IDF, Machine Learning, SVM

Abstrak

Perkembangan perdagangan elektronik telah mendorong peningkatan transaksi produk *Muslim Fashion* sekaligus menghasilkan sejumlah besar ulasan konsumen yang mengandung informasi mengenai pengalaman dan persepsi terhadap produk. Analisis terhadap ulasan tersebut penting dilakukan karena informasi sentimen dapat membantu memahami kecenderungan kepuasan konsumen secara lebih sistematis dibandingkan analisis manual. Penelitian ini bertujuan mengklasifikasikan sentimen ulasan konsumen terhadap produk *Muslim Fashion* serta membandingkan kinerja beberapa algoritma *machine learning* dalam melakukan klasifikasi sentimen. Dataset penelitian diperoleh dari PRDECT-ID yang berjumlah 5.400 ulasan. Berdasarkan proses penyaringan kategori *Muslim Fashion*, diperoleh 200 ulasan. Selanjutnya, sentimen ditentukan berdasarkan isi *Customer Review*, kemudian data dengan label *Neutral* dikeluarkan sehingga diperoleh 135 ulasan, terdiri atas 97 ulasan negatif dan 38 ulasan positif. Tahapan penelitian meliputi *preprocessing* teks, pembobotan fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF), klasifikasi menggunakan Multinomial Naive Bayes (MNB), Logistic Regression (LR), dan Support Vector Machine (SVM), serta evaluasi menggunakan Stratified 5-Fold Cross Validation. Kontribusi penelitian ini adalah memberikan perbandingan empiris ketiga algoritma pada ulasan produk *Muslim Fashion* dengan mempertimbangkan ketidakseimbangan kelas dan beberapa metrik evaluasi. Hasil eksperimen menunjukkan bahwa SVM memberikan performa terbaik dengan Accuracy sebesar 84,44%, Precision 81,43%, Recall 60,36%, dan F1-Score 68,00%. Hasil tersebut menunjukkan bahwa kombinasi TF-IDF dan SVM lebih efektif dibandingkan MNB dan LR dalam klasifikasi sentimen pada dataset penelitian.

Kata Kunci: Muslim Fashion, Klasifikasi Sentimen, TF-IDF, Machine Learning, SVM

<https://journal.thamrin.ac.id/index.php/jtik/article/view/3777/3042>

PENDAHULUAN

Perkembangan industri fashion Muslim di Indonesia menunjukkan adanya perubahan pola konsumsi masyarakat yang semakin terintegrasi dengan perdagangan elektronik. Produk fashion Muslim tidak hanya berkaitan dengan pemenuhan kebutuhan berpakaian, tetapi juga berhubungan dengan identitas, gaya hidup, nilai keagamaan, dan preferensi konsumen Muslim. Penelitian (Ayumadani & Zahra, 2025) menunjukkan bahwa konsumsi halal fashion di Indonesia berkaitan dengan pengetahuan konsumen, kesadaran halal, gaya hidup, dan keputusan pembelian. Penelitian tersebut juga menunjukkan adanya perbedaan karakteristik perilaku antar-generasi, sehingga pemahaman terhadap perilaku dan respons konsumen menjadi penting bagi pengembangan industri fashion Muslim.

Di sisi lain, perkembangan perdagangan elektronik menghasilkan sejumlah besar ulasan konsumen yang mengandung informasi mengenai pengalaman, kepuasan, ketidakpuasan, kualitas produk, harga, pelayanan, pengiriman, dan karakteristik produk. Informasi tersebut merupakan sumber data yang berharga untuk memahami persepsi konsumen, tetapi jumlah data yang besar menyebabkan analisis secara manual menjadi tidak efisien. Natural Language Processing (NLP) dan machine learning dapat digunakan untuk mengubah data ulasan yang tidak terstruktur menjadi informasi sentimen yang dapat dianalisis secara sistematis.

Urgensi penelitian ini terletak pada kebutuhan untuk memperoleh gambaran berbasis data mengenai kecenderungan sentimen konsumen terhadap produk Muslim Fashion. Analisis sentimen dapat membantu mengidentifikasi apakah pengalaman konsumen cenderung positif atau negatif serta memberikan informasi awal mengenai penerimaan konsumen terhadap produk yang dipasarkan melalui perdagangan elektronik. Selain itu, penggunaan dataset PRDECT-ID yang menyediakan ulasan produk berbahasa Indonesia memungkinkan penelitian dilakukan pada data konsumen Indonesia dengan pendekatan yang dapat direproduksi. Dataset tersebut terdiri atas 5.400 ulasan dari berbagai kategori produk dan memiliki label sentimen positif dan negatif. Dengan demikian, penelitian ini penting dilakukan untuk menguji sejauh mana metode machine learning berbasis representasi teks TF-IDF dapat digunakan untuk mengklasifikasikan sentimen konsumen pada kategori Muslim Fashion. Hasil penelitian diharapkan dapat memberikan informasi

empiris mengenai performa algoritma machine learning dalam mengolah ulasan konsumen berbahasa Indonesia pada domain fashion Muslim.

Penelitian mengenai fashion Muslim di Indonesia telah dilakukan dari berbagai perspektif. (Zahro et al., 2025) misalnya, menganalisis perilaku pembelian halal fashion menggunakan pendekatan survei terhadap 504 konsumen Muslim dan metode Partial Least Squares Multi-Group Analysis (PLS-MGA). Penelitian tersebut berfokus pada hubungan antara pengetahuan konsumen, kesadaran halal, gaya hidup, dan keputusan pembelian, bukan pada pengolahan data tekstual berupa ulasan konsumen.

Penelitian yang menganalisis ulasan daring pada sektor Muslim women's fashion (Nurchahyanie & Saraswati, 2026). Penelitian tersebut menggunakan 2.050 online product reviews dari tiga brand Muslim fashion di Indonesia dan mengembangkan pendekatan dua lapis yang menggabungkan klasifikasi sentimen dengan interpretasi semantik berbasis ahli untuk menerjemahkan hasil sentimen menjadi atribut desain produk.

Penelitian ini tidak diarahkan untuk menentukan status halal suatu produk, tetapi untuk mengidentifikasi kecenderungan sentimen konsumen terhadap produk yang berada pada kategori Muslim Fashion. Selain itu, penelitian ini memberikan perbandingan langsung beberapa algoritma machine learning berbasis TF-IDF pada domain tersebut sehingga dapat diketahui algoritma yang paling sesuai untuk klasifikasi ulasan konsumen berbahasa Indonesia.

Dengan demikian, research gap penelitian ini terletak pada perbedaan objek, pendekatan data, dan fokus analisis dibandingkan penelitian terdahulu. Penelitian sebelumnya lebih banyak menggunakan survei untuk mengkaji perilaku pembelian halal fashion atau menggunakan kumpulan ulasan Muslim fashion yang lebih besar dengan pendekatan hybrid dan interpretasi semantik. Penelitian ini berfokus pada eksperimen klasifikasi sentimen berbasis machine learning klasik terhadap subset Muslim Fashion pada PRDECT-ID dengan prosedur preprocessing dan evaluasi yang terstruktur.

Penelitian ini bertujuan untuk 1) Mengidentifikasi dan menganalisis karakteristik sentimen konsumen pada ulasan produk yang termasuk dalam kategori Muslim Fashion berdasarkan dataset PRDECT-ID. 2) Mengembangkan model klasifikasi sentimen berbasis machine learning menggunakan representasi fitur Term Frequency-Inverse Document

Frequency (TF-IDF) untuk mengklasifikasikan ulasan konsumen ke dalam sentimen positif dan negatif. 3) Membandingkan performa algoritma Multinomial Naive Bayes, Logistic Regression, dan Support Vector Machine (SVM) berdasarkan accuracy, precision, recall, dan F1-score untuk menentukan algoritma yang memberikan kinerja terbaik dalam klasifikasi sentimen konsumen terhadap produk Muslim Fashion. Penelitian ini memberikan beberapa kontribusi, baik secara akademis maupun praktis. Kontribusi pada penelitian fashion Muslim, penelitian ini memberikan perspektif berbasis data mengenai kecenderungan sentimen konsumen yang diperoleh dari ulasan produk. Berbeda dengan penelitian yang menggunakan survei untuk mengukur faktor-faktor yang memengaruhi keputusan pembelian halal fashion, penelitian ini memanfaatkan data tekstual aktual yang berasal dari ulasan konsumen. Pendekatan ini dapat melengkapi penelitian perilaku konsumen dengan memberikan gambaran mengenai pengalaman dan persepsi konsumen yang diekspresikan secara langsung dalam ulasan.

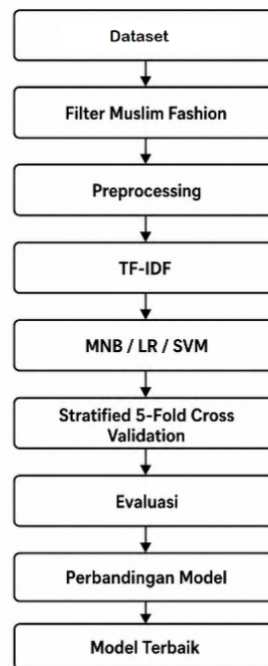
Penelitian ini menghasilkan perbandingan performa Multinomial Naive Bayes, Logistic Regression, dan Support Vector Machine pada subset Muslim Fashion. Hasil perbandingan dapat menunjukkan algoritma yang paling sesuai untuk klasifikasi sentimen pada dataset berbahasa Indonesia dengan ukuran data yang relatif terbatas.

Secara keseluruhan, kontribusi utama penelitian ini adalah menyediakan eksperimen yang terukur dan dapat direproduksi untuk menganalisis sentimen konsumen pada domain Muslim Fashion menggunakan machine learning berbasis TF-IDF, sekaligus memberikan bukti empiris mengenai algoritma yang paling efektif pada data ulasan berbahasa Indonesia.

METODE PENELITIAN

Metode yang Diusulkan

Langkah penelitian pada studi ini diawali dengan 1) pengumpulan dataset PRDECT-ID 2) filtering data berdasarkan kategori Muslim Fashion 3) preprocessing teks 4) Ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) 5) klasifikasi Multinomial Naive Bayes (MNB), Logistic Regression (LR), dan Support Vector Machine (SVM). 6) evaluasi menggunakan Stratified 5-Fold Cross Validation 7) evaluasi. Keseluruhan tahapan penelitian tersebut disajikan pada Gambar 1.



Gambar 1. Langkah Penelitian

Term Frequency-Inverse Document Frequency (TF-IDF)

TF-IDF merupakan metode pembobotan term yang banyak digunakan dalam representasi teks untuk tugas klasifikasi, dengan memberikan bobot berdasarkan frekuensi kemunculan term dalam dokumen dan tingkat kemunculannya pada keseluruhan korpus (Jiang et al., 2021). Dalam Penelitian lain, Algoritma Term Frequency-Inverse Document Frequency (TF-IDF) merupakan skema pembobotan istilah paling populer, digunakan untuk menilai relevansi suatu kata dalam dokumen secara numerik. Skor frekuensi yang diberikan pada suatu kata melalui TF-IDF menentukan tingkat kepentingan kata tersebut bagi dokumen berdasarkan frekuensi kemunculannya (Addiga & Bagui, 2022). Rumus yang digunakan untuk menghitung skor TF-IDF secara bertahap adalah sebagai berikut:

$$TF(t, d) = \frac{f_{t,d}}{\sum_k f_{k,d}}$$

Keterangan:

- t = term/kata
- d = dokumen
- $f_{t,d}$ = jumlah kemunculan term t pada dokumen d
- $\sum_k f_{k,d}$ = jumlah seluruh term dalam dokumen d

Multinomial Naive Bayes

Multinomial Naive Bayes adalah model pembelajaran probabilistik yang diturunkan dari teori Bayes. Model ini banyak digunakan dalam bidang Pemrosesan Bahasa Alami (NLP) dan beroperasi berdasarkan prinsip frekuensi istilah (Term Frequency), yang mengukur seberapa sering suatu kata muncul dalam sebuah dokumen. Model ini memberikan penjelasan mengenai dua aspek: keberadaan atau ketidakberadaan suatu kata dalam teks, serta frekuensi kemunculan kata tersebut di dalam dokumen. Multinomial Naive Bayes merupakan model klasifikasi teks yang memanfaatkan pembelajaran terawasi (Supervised Learning). Oleh karena itu, model ini memerlukan data yang telah diberi label sebelum proses pelatihan dapat dimulai. Formulasi Multinomial Naive Bayes dapat dinyatakan melalui persamaan (Pratama et al., 2024).

Logistic Regression

Logistic Regression merupakan jenis regresi analisis yang menjelaskan hubungan antara variabel dependen dan variabel independen. Ini menjelaskan hubungan antara satu atau lebih variabel bebas dan variabel terikat dari jenis kategori tertentu, yang dapat berkisar antara 0 dan 1 dan benar atau salah, besar atau kecil (Putri et al., 2025).

Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma klasifikasi tangguh yang mampu melakukan klasifikasi baik secara linear maupun non-linear. Gagasan utama SVM adalah menemukan hyperplane optimal yang memaksimalkan margin antara dua kelas, sehingga memastikan titik-titik data terdekat yang disebut sebagai support vector berada sejauh mungkin dari batas keputusan. Ketika data tidak dapat dipisahkan secara linear, SVM menggunakan fungsi kernel untuk memetakan data ke dalam ruang berdimensi lebih tinggi, yang memungkinkan pemisahan pola-pola kompleks secara lebih efektif. Dengan memaksimalkan margin geometris antar-kelas, SVM memiliki kemampuan generalisasi yang kuat, menjadikannya salah satu algoritma paling andal untuk tugas analisis sentimen (Lestari & Hutagalung, 2025).

Dataset

Dataset pada penelitian ini menggunakan PRDECT-ID (Product Review Dataset for Emotion and Sentiment Analysis in Indonesian) yang terdiri atas 5.400 ulasan konsumen

berbahasa Indonesia dari berbagai kategori produk. Dataset tersebut digunakan sebagai sumber data penelitian karena menyediakan informasi ulasan konsumen beserta label sentimen yang dapat digunakan untuk proses klasifikasi.

HASIL DAN PEMBAHASAN

Dataset

Dataset yang digunakan dalam penelitian ini adalah PRDECT-ID (Product Review Dataset for Emotion and Sentiment Analysis in Indonesian) yang terdiri atas 5.400 ulasan konsumen berbahasa Indonesia dari berbagai kategori produk. Dataset tersebut menyediakan informasi ulasan konsumen dan label sentimen yang digunakan sebagai kelas target dalam penelitian.

Untuk menyesuaikan dataset dengan fokus penelitian, dilakukan proses filtering berdasarkan atribut category dengan memilih kategori Muslim Fashion. Proses tersebut menghasilkan 200 ulasan yang selanjutnya digunakan sebagai dataset penelitian. Contoh data berdasar kategori dapat dilihat di Tabel 1.

Tabel 1. Dataset

Category	Product Name	Customer Review
Women's Fashion	COMFEET - Kaos Kaki Mata Kaki / Ankle Socks Import Quality - White	oke sih ini. bahannya enak buat dipakai. tebal tipisnya jg perfect buat sy, ga ketebelan ga ketipisan
Women's Fashion	COMFEET - Kaos Kaki Mata Kaki / Ankle Socks Import Quality - White	Bahannya adem. Tebalnya pas!. Jahitan rapi. Sudah lngganan beli d sini dan selalu repeat order.....recomended seeller.....
Women's Fashion	COMFEET - Kaos Kaki Mata Kaki / Ankle Socks Import Quality - White	Pengiriman cepat. barang lumayan
Women's Fashion	COMFEET - Kaos Kaki Mata Kaki / Ankle Socks Import Quality - White	Bahannya adem Warnanya cantik ?? Tebalnya pas! Jahitan rapi
Women's Fashion	COMFEET - Kaos Kaki Mata Kaki / Ankle Socks Import Quality - White	Barang bagus, sesuai kualitas dan harga. Proses dan respon agak lambat mesti diingatkan dulu
Women's Fashion	COMFEET - Kaos Kaki Mata Kaki / Ankle Socks Import Quality - White	Produk bgs packing bgs pengiriman cpt
Women's Fashion	COMFEET - Kaos Kaki Mata Kaki / Ankle Socks Import Quality - White	Produknya sesuai gambar

Women's Fashion	COMFEET - Kaos Kaki Mata Kaki / Ankle Socks Import Quality - White	Pengirimannya cepat, barang bagus, tapi ukuran maksimal 27 sepertinya, jadi tidak muat
Women's Fashion	KK023 Kaos Kaki Pendek Japanese Cool Solid Line Wanita - GINGER	bahannya enak gak gerah dipakainya.
Women's Fashion	KK023 Kaos Kaki Pendek Japanese Cool Solid Line Wanita - GINGER	beli buat ibu saya, muat sampai ukuran besar. waktu dikomplain ada yg cacat, seller cepet respon & diganti dgn yg baru. terimakasih ??
Women's Fashion	KK023 Kaos Kaki Pendek Japanese Cool Solid Line Wanita - GINGER	Puas banget kaos kakinya bagus
Women's Fashion	KK023 Kaos Kaki Pendek Japanese Cool Solid Line Wanita - GINGER	Warnanya cantik ?? Bahannya adem Tebalnya pas! Jahitan rapi. Ini keren banget si harga terjangkau kualitasnya oke.. pelayanannya okee
Women's Fashion	KK023 Kaos Kaki Pendek Japanese Cool Solid Line Wanita - GINGER	Warna sesuai, ada karet di bagian belakang spy kaos kaki tdk mudah lepas ????
Women's Fashion	KK023 Kaos Kaki Pendek Japanese Cool Solid Line Wanita - GINGER	sesuai pesanan trmkasi.
.... N: 200		

Sebanyak 200 ulasan pada kategori Muslim Fashion selanjutnya dianalisis berdasarkan isi *Customer Review* untuk menentukan kecenderungan sentimennya. Analisis dilakukan menggunakan model IndoBERT, yang memproses setiap ulasan dan menghasilkan label sentimen berupa *Positive*, *Negative*, atau *Neutral* beserta nilai *confidence score*. Hasil analisa dapat dilihat dalam Tabel 2.

Tabel 2. Dataset Sentiment

Category	Product Name	Customer Review	Sentiment	Confidence
Muslim Fashion	HIJAB SEGI EMPAT AUTHENTISM BASIC RACHITA VOAL/LAMINA VOAL / ULTRAFINE - TALYA	ternyata tipis bahannya	LABEL_0	0.995634
Muslim Fashion	HIJAB SEGI EMPAT AUTHENTISM BASIC RACHITA VOAL/LAMINA VOAL / ULTRAFINE - TALYA	deskripsi warna kurang sesuai	LABEL_1	0.999277
Muslim Fashion	Hijab VOAL Segiempat Premium - Emikoawa Jilbab Kerudung Terbaru Korea - capucino	kurang bagus bahannya	LABEL_1	0.997913

Muslim Fashion	masker hijab-masker hijab serut-masker kain-masker hijab 2ply 2 lapis - Ungu tua, Merk Mambo	gk bisa di pakai tali krng panjang	LABEL_1	0.997741
Muslim Fashion	masker hijab-masker hijab serut-masker kain-masker hijab 2ply 2 lapis - Ungu tua, Merk Mambo	sangat kecewa belanja di toko ini pesan 12 masker yang sampai cuma 10 tolong kalau packing barang sbelum di kirim teliti dlu dong, jangan sampai mengecewakan pembeli termasuk saya .	LABEL_2	0.838901
Muslim Fashion	Masker Hijab 3 ply Masker Headloop Kain Polos Warna Pastel model Evo - Kubus	barang g smapai hilang di kurir ..uang refund	LABEL_2	0.999707
Muslim Fashion	Masker Hijab 3 ply Masker Headloop Kain Polos Warna Pastel model Evo - Kubus	Pesan lain warna dikirim. Warna yg sama. Kecewa seh tp ya sdhlah	LABEL_1	0.998098
Muslim Fashion	Masker Hijab 3 ply Masker Headloop Kain Polos Warna Pastel model Evo - Kubus	Penjual tidak membaca catatan yang diberikan pembeli	LABEL_2	0.999435
... 200				

Proses selanjutnya adalah melakukan mapping label karena hasil dari proses sebelumnya masih menggunakan Label_0, Label_2 dan Label_3. Mapping label yang diharapkan adalah sentiment positive dan negative. Hasil Mapping dapat dilihat dalam Tabel 3.

Tabel 3. Hasil Mapping Sentiment

Sentiment	Jumlah Data
positive	97
negative	38
neutral	65
Jumlah	200

Sentiment yang digunakan dalam penelitian adalah positive dan negative sehingga sentiment neutral tidak dipakai. Maka data yang memenuhi syarat adalah 135. Dari data 135 selanjutnya dicari prosentase sentiment. Hasil perhitungan dapat dilihat dalam tabel 4.

Tabel 4. Hasil Presentase Sentiment

Sentiment	Persentase
positive	28.15
negative	71.85

Tahap berikutnya adalah preprocessing teks. Tahapan ini bertujuan mengubah teks ulasan yang masih tidak terstruktur menjadi data yang lebih sesuai untuk proses klasifikasi. Preprocessing yang digunakan terdiri atas 1) *Case folding* 2) *Cleaning* 3) Tokenisasi 4) Normalisasi 5) *Stopword removal* 6) *Stemming*. Case Folding: Seluruh huruf diubah menjadi huruf kecil. Tujuannya agar kata Bagus, BAGUS, dan bagus dianggap sebagai term yang sama. Cleaning: Karakter yang tidak diperlukan seperti tanda baca, angka tertentu, URL, emoji, dan karakter khusus dihilangkan. Tokenisasi: Kalimat kemudian dipecah menjadi token atau kata. Contoh: [barangnya, bagus, banget, bahannya, nyaman]. Normalisasi: Kata tidak baku atau bentuk singkatan dapat diubah menjadi bentuk yang lebih baku. Contoh: bgt menjadi banget. Stopword Removal: Kata-kata yang memiliki kontribusi rendah terhadap proses klasifikasi dapat dihilangkan. Contoh: yang, dan, di, ke, dari, ini, itu. Sehingga didapatkan barangnya menjadi barang. Hasil akhir preprocessing inilah yang digunakan pada tahap ekstraksi fitur. Hasil proses preprocessing dapat dilihat dalam Tabel 5.

Tabel 5. Hasil Preprocessing

Customer Review	Processed Review
ternyata tipis bahannya	nyata tipis bahan
sangat kecewa belanja di toko ini pesan 12 masker yang sampai cuma 10 tolong kalau packing barang sbelum di kirim teliti dlu dong, jangan sampai mengecewakan pembeli termasuk saya .	sangat kecewa belanja toko pesan masker cuma kalau packing barang sbelum kirim teliti dlu dong jangan sampai kecewa beli masuk
barang g smapai hilang di kurir ..uang refund	barang g smapai hilang kurir uang refund
Penjual tidak membaca catatan yang diberikan pembeli	jual baca catat beri beli
Minta 4 pcs dikirimnya 3 pcs dan dikemas paket tidak ada jumlah barang yang akan dikirim.	minta pcs kirim pcs kemas paket ada jumlah barang akan kirim
gasuka warnanya:(	gasuka warna
saya nggak terima paketnya sama sekali saya minta lacak malah no respon	nggak terima paket sama sekali minta lacak malah no respon
order emerald, dikirim tosca tanpa konfirmasi. ??	order emerald kirim tosca konfirmasi
tipis.kecil	tipis kecil
...	

Term Frequency-Inverse Document Frequency (TV-IDF)

Setelah melalui tahap preprocessing, teks ulasan dikonversi menjadi representasi numerik menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Metode TF-IDF digunakan untuk memberikan bobot terhadap setiap term berdasarkan tingkat kemunculannya dalam suatu dokumen dan tingkat kepentingannya dibandingkan dengan keseluruhan dokumen. Term yang sering muncul pada suatu dokumen tetapi relatif jarang ditemukan pada dokumen lainnya akan memperoleh bobot yang lebih tinggi.

Tahap pertama adalah menghitung Term Frequency (TF), yaitu frekuensi kemunculan suatu term dalam sebuah dokumen. Pada penelitian ini, perhitungan TF dilakukan menggunakan *script* Python. Hasil proses ini dapat ditunjukkan dalam Tabel 6.

Tabel 6. Hasil TF

Document	Sentiment	aamiin	abu	ada	adeeeeem	adem	admin	aja	aju	akan	...
1	Negative	0	0	0	0	0	0	0	0	0	...
2	Positive	0	0	0	0	0	0	0	0	0	
3	Positive	0	0	0	0	0	0	0	0	0	
4	Positive	0	0	0	0	0	0	0	0	0	
5	Positive	0	0	1	0	0	0	0	0	1	
6	Negative	0	0	0	0	0	0	0	0	0	
7	Positive	0	0	0	0	0	0	0	0	0	
8	Negative	0	0	0	0	0	0	1	0	0	
9	Positive	0	0	0	0	0	0	1	0	0	
10	Positive	0	0	0	0	0	0	0	0	0	
...											

Tahap selanjutnya adalah menghitung Document Frequency adalah berapa banyak dokumen yang mengandung setiap term. Hasil perhitungan dapat dilihat dalam Tabel 7.

Tabel 7. Hasil DF

Term	Document Frequency
warna	46
bahan	44
sesuai	38
bagus	31
cantik	29
kirim	26
yg	13
mudah	11
...	

Berdasar data DF, tahap selanjutnya adalah menghitung Inverse Document Frequency (IDF). Hasil IDF dapat dilihat dalam Tabel 8.

Tabel 8. Hasil IDF

Term	Document Frequency	IDF
warna	46	1.076633
bahan	44	1.121085
sesuai	38	1.267689
bagus	31	1.471288
cantik	29	1.537979
rapi	27	1.609438
...		

Menghitung TF-IDF. Hasil Perhitungan dapat dilihat dalam tabel 9.

Tabel 9. Hasil TD-IDF

Document	Sentiment	aamiin	abu	ada	adeeeeem	adem	admin	aja	aju	akan	aku	...
1	Negative	0	0	0	0	0	0	0	0	0	0	
2	Positive	0	0	0	0	0	0	0	0	0	0	
3	Positive	0	0	0	0	0	0	0	0	0	0	
4	Positive	0	0	0	0	0	0	0	0	0	0	
5	Positive	0	0	0.32	0	0	0	0	0	0.44	0	
6	Negative	0	0	0	0	0	0	0	0	0	0	
7	Positive	0	0	0	0	0	0	0	0	0	0	
...												

Klasifikasi MNB, LR, dan SVM

Tahap klasifikasi dilakukan menggunakan tiga algoritma *machine learning*, yaitu Multinomial Naive Bayes (MNB), Logistic Regression (LR), dan Support Vector Machine (SVM). Teks hasil preprocessing digunakan sebagai masukan, sedangkan label *Positive* dan *Negative* digunakan sebagai target klasifikasi. Sebelum proses klasifikasi, teks ditransformasikan ke dalam representasi TF-IDF menggunakan *pipeline* sehingga proses pembentukan bobot TF-IDF dilakukan secara terpisah pada data pelatihan pada setiap *fold*. Pendekatan tersebut diterapkan untuk mencegah *data leakage* antara data pelatihan dan data pengujian. Setiap algoritma dievaluasi menggunakan Stratified 5-Fold Cross Validation dengan *shuffle* dan *random state* 42. Evaluasi dilakukan menggunakan Accuracy, Precision, Recall, dan F1-Score. Nilai performa akhir setiap algoritma diperoleh dari rata-rata hasil

pada lima *fold*. Algoritma dengan nilai F1-Score rata-rata tertinggi ditetapkan sebagai model dengan performa terbaik. Hasil dari perhitungan Klasifikasi dapat dilihat dalam tabel 10.

Tabel 10. Hasil Klasifikasi

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Std Accuracy (%)	Std F1-Score (%)
Multinomial Naive Bayes	76.3	73.33	18.21	28.61	2.96	16.12
Logistic Regression	75.56	80	13.21	22.22	2.96	14.05
Support Vector Machine	84.44	81.43	60.36	68	3.63	8.36

Stratified 5-Fold Cross Validation

Stratified 5-Fold Cross Validation digunakan untuk memperoleh estimasi performa yang lebih stabil. Dataset dibagi menjadi lima bagian (*fold*) dengan tetap mempertahankan proporsi masing-masing kelas sentimen. Pada setiap iterasi, empat *fold* digunakan sebagai data pelatihan dan satu *fold* digunakan sebagai data pengujian. Proses tersebut diulang sebanyak lima kali sehingga setiap data memperoleh kesempatan menjadi data pengujian. Nilai performa akhir setiap algoritma diperoleh dari rata-rata hasil kelima *fold*. Hasil perhitungan dapat dilihat dalam Tabel 11.

Tabel 11. Hasil Stratified 5-Fold Cross Validation

Model	Fold	Accuracy	Precision	Recall	F1-Score
Multinomial Naive Bayes	1	0.7407	0	0	0
Multinomial Naive Bayes	2	0.8148	1	0.2857	0.4444
Multinomial Naive Bayes	3	0.7778	1	0.25	0.4
Multinomial Naive Bayes	4	0.7407	1	0.125	0.2222
Multinomial Naive Bayes	5	0.7407	0.6667	0.25	0.3636
Logistic Regression	1	0.7407	0	0	0
Logistic Regression	2	0.8148	1	0.2857	0.4444
Logistic Regression	3	0.7407	1	0.125	0.2222
Logistic Regression	4	0.7407	1	0.125	0.2222
Logistic Regression	5	0.7407	1	0.125	0.2222
Support Vector Machine	1	0.8519	1	0.4286	0.6
Support Vector Machine	2	0.8519	0.7143	0.7143	0.7143
Support Vector Machine	3	0.8519	0.8333	0.625	0.7143
Support Vector Machine	4	0.7778	0.6667	0.5	0.5714
Support Vector Machine	5	0.8889	0.8571	0.75	0.8

Berdasarkan hasil klasifikasi terhadap 135 ulasan produk Muslim Fashion, diketahui bahwa dataset terdiri atas 97 ulasan negatif dan 38 ulasan positif. Distribusi tersebut menunjukkan adanya ketidakseimbangan kelas, dengan proporsi kelas negatif sebesar 71,85% dan kelas positif sebesar 28,15%. Kondisi ketidakseimbangan kelas perlu diperhatikan dalam interpretasi hasil evaluasi karena model dapat memperoleh nilai Accuracy yang relatif tinggi dengan lebih sering memprediksi kelas mayoritas.

Hasil Stratified 5-Fold Cross Validation menunjukkan bahwa Multinomial Naive Bayes memperoleh nilai Accuracy rata-rata sebesar 76,30%, Precision 73,33%, Recall 18,21%, dan F1-Score 28,61%. Meskipun Accuracy MNB mencapai 76,30%, nilai Recall yang rendah menunjukkan bahwa model memiliki keterbatasan dalam mengidentifikasi ulasan positif. Hal tersebut terlihat pada Fold 1, ketika model memperoleh Recall sebesar 0% dan F1-Score sebesar 0%, yang menunjukkan bahwa tidak terdapat prediksi positif yang benar pada fold tersebut.

Logistic Regression memperoleh Accuracy rata-rata sebesar 75,56%, Precision 80,00%, Recall 13,21%, dan F1-Score 22,22%. Nilai Precision yang relatif tinggi menunjukkan bahwa ketika model memprediksi suatu ulasan sebagai positif, prediksi tersebut relatif tepat. Namun, nilai Recall yang hanya 13,21% menunjukkan bahwa sebagian besar ulasan positif belum berhasil dikenali oleh model. Kondisi tersebut menghasilkan F1-Score yang rendah.

Berbeda dengan kedua algoritma tersebut, Support Vector Machine memberikan performa yang lebih baik dengan Accuracy sebesar 84,44%, Precision 81,43%, Recall 60,36%, dan F1-Score 68,00%. SVM tidak hanya menghasilkan Accuracy tertinggi, tetapi juga menunjukkan kemampuan yang lebih baik dalam mengenali kelas positif. Nilai Recall SVM yang mencapai 60,36% jauh lebih tinggi dibandingkan MNB sebesar 18,21% dan Logistic Regression sebesar 13,21%.

KESIMPULAN DAN REKOMENDASI

Berdasarkan hasil penelitian Analisis Sentimen Konsumen terhadap Produk Muslim Fashion Menggunakan Machine Learning, dapat disimpulkan bahwa dataset PRDECT-ID sebanyak 5.400 ulasan setelah dilakukan penyaringan berdasarkan kategori Muslim Fashion

menghasilkan 200 ulasan. Selanjutnya, analisis sentimen berdasarkan Customer Review menghasilkan tiga kelas, yaitu Positive, Negative, dan Neutral. Karena penelitian menggunakan klasifikasi biner, ulasan dengan sentimen Neutral dikeluarkan sehingga diperoleh 135 data, yang terdiri atas 97 ulasan negatif dan 38 ulasan positif.

Representasi teks menggunakan Term Frequency-Inverse Document Frequency (TF-IDF) berhasil mengubah data ulasan menjadi representasi numerik yang dapat digunakan sebagai masukan bagi algoritma klasifikasi. Selanjutnya, penelitian membandingkan tiga algoritma, yaitu Multinomial Naive Bayes (MNB), Logistic Regression (LR), dan Support Vector Machine (SVM) menggunakan Stratified 5-Fold Cross Validation.

Hasil eksperimen menunjukkan bahwa Support Vector Machine merupakan algoritma dengan performa terbaik dengan nilai Accuracy sebesar 84,44%, Precision 81,43%, Recall 60,36%, dan F1-Score 68,00%. Sementara itu, MNB memperoleh Accuracy 76,30% dan F1-Score 28,61%, sedangkan LR memperoleh Accuracy 75,56% dan F1-Score 22,22%. Hasil tersebut menunjukkan bahwa SVM memiliki kemampuan yang lebih baik dalam mengenali sentimen dibandingkan MNB dan LR pada dataset penelitian.

REFERENSI

- Addiga, A., & Bagui, S. (2022). Sentiment analysis on twitter data using term frequency-inverse document frequency. *Journal of Computer and Communications*, 10(8), 117–128.
- Ayumadani, P., & Zahra, I. A. (2025). Pengaruh Halal Awareness dan Celebrity Endorsement terhadap Keputusan Pembelian Skincare Halal (Studi Kasus Gen Z di Jakarta). *Seminar Nasional Inovasi Vokasi*, 4, 1036–1043.
- Aprianti, Y, Hananto, AL, & Hilabi, SS (2025). Klasifikasi sentimen komentar pengguna pada aplikasi Ruangguru menggunakan algoritma Naive Bayes. *Metik Jurnal*, journal.universitasmulia.ac.id, <https://journal.universitasmulia.ac.id/index.php/metik/article/view/1023>
- Berutu, SS (2022). Text Mining dan Klasifikasi Sentimen Berbasis Naïve Bayes Pada Opini Masyarakat terhadap Makanan Tradisional. ... *Sistem Komputer dan Informatika (JSON)* Hal, pdfs.semanticscholar.org, <https://journal.thamrin.ac.id/index.php/jtik/article/view/3777/3042>

<https://pdfs.semanticscholar.org/15dc/e5cda97fd7fad4446a00a7500149aa4e77c9.pdf>
f

- Fitroh, QA, Mardlian, MSAK, & ... (2026). Evaluasi Strategi Penanganan Ketidakseimbangan Kelas pada Klasifikasi Sentimen Multikelas Ulasan Aplikasi ChatGPT. *JUSIFOR: Jurnal ...*, ejournal.uniramalang.ac.id, <https://ejournal.uniramalang.ac.id/jusifor/article/view/9714>
- Irawan, B, & Bachri, OS (2026). Klasifikasi Sentimen Ulasan Produk Samsung pada Platform E-Commerce Tokopedia Menggunakan Algoritma Naive Bayes. *Rabit: Jurnal Teknologi dan Sistem ...*, jurnal.univrab.ac.id, <https://jurnal.univrab.ac.id/index.php/rabit/article/view/7557>
- Jiang, Z., Gao, B., He, Y., Han, Y., Doyle, P., & Zhu, Q. (2021). Text Classification Using Novel Term Weighting Scheme-Based Improved TF-IDF for Internet Media Reports. *Mathematical Problems in Engineering*, 2021(1), 6619088.
- Lestari, V. B., & Hutagalung, C. A. (2025). Evaluation of TF-IDF Extraction Techniques in Sentiment Analysis of Indonesian-Language Marketplaces Using SVM, Logistic Regression, and Naive Bayes. *J-KOMA: Jurnal Ilmu Komputer Dan Aplikasi*, 8(1), 36–44.
- Munandar, AA, Farikhin, F, & Widodo, CE (2023). Sentimen Analisis Aplikasi Belajar Online Menggunakan Klasifikasi SVM. *Comput. Sci., vol. 8, no. 2, p ...*, academia.edu, <https://www.academia.edu/download/123663867/2563.pdf>
- Mubarak, MZ, Rizqi, S, Rizal, RM, & ... (2025). Klasifikasi Sentimen Ulasan Aplikasi Claude Menggunakan Naive Bayes dan Support Vector Machine. *JUSIFOR: Jurnal ...*, ejournal.uniramalang.ac.id, <https://ejournal.uniramalang.ac.id/jusifor/article/view/7054>
- Nurcahyanie, Y. D., & Saraswati, S. N. (2026). Text-Based Sentiment Analysis of Online Reviews: Evidence from Indonesia's Muslim Women's Fashion Sector. *Bit-Tech*, 8(3), 4144–4156.
- Pratama, D. R. S., Munandar, T. A., & Ramdhania, K. F. (2024). Multinomial naive bayes algorithm for indonesian language sentiment classification related to jakarta

-
- international stadium (jis). *International Journal of Information Technology and Computer Science Applications*, 2(1), 12–22.
- Putri, A. A., Agustian, S., Abdillah, R., & Pizaini, P. (2025). Penerapan Metode Logistic Regression Untuk Klasifikasi Sentimen Pada Dataset Twitter Terbatas. *ZONasi: Jurnal Sistem Informasi*, 7(1), 95–107.
- Siswoyo, B., & Utomo, NAP (2025). Pemanfaatan machine learning untuk klasifikasi sentimen pelanggan pada media sosial. *Jurnal Teknologi Informasi Digital*, jurnal.ipdig.id, <https://jurnal.ipdig.id/index.php/jtid/article/view/22>
- Zahro, Z., Affa, A. S., Jannah, K. D., Shofiyuddin, M., Irawan, Z., & Fahmi, L. N. (2025). From knowledge to decision: A PLS-MGA investigation of halal fashion purchase behaviour across generation in Indonesia. *Diponegoro International Journal of Business*, 8(2), 107–128.