

PENERAPAN LEXICON-BASED LABELING UNTUK ANALISIS SENTIMEN ULASAN UMKM KULINER DENGAN PERBANDINGAN KINERJA ALGORITMA MACHINE LEARNING

Siti Rihastuti¹, Afnan Rosyidi²

^{1,2}STMIK AMIKOM Surakarta, Sukoharjo

¹siti@dosen.amikomsolo.ac.id

²afnan@dosen.amikomsolo.ac.id

Abstrak

Penelitian ini bertujuan mengklasifikasikan sentimen konsumen UMKM kuliner di Kalimantan Selatan menggunakan pendekatan hybrid yang menggabungkan metode Lexicon-Based dan algoritma Machine Learning. Dataset yang digunakan berupa kumpulan ulasan konsumen dari platform Instagram dan TikTok. Lexicon-based digunakan untuk pelabelan otomatis berdasarkan kamus kata positif dan negatif, sementara enam algoritma Machine Learning yaitu Naive Bayes, Random Forest, Logistic Regression, Support Vector Machine (SVM), Decision Tree, dan K-Nearest Neighbors (KNN) digunakan sebagai model klasifikasi sentimen berbasis pola. Tahapan penelitian mencakup pengambilan dataset, text preprocessing, pelabelan otomatis, ekstraksi fitur TF-IDF, penerapan model, dan evaluasi menggunakan metrik akurasi, precision, recall, serta F1-score. Hasil penelitian menunjukkan bahwa SVM unggul dengan akurasi 79% dan F1-score 70%, diikuti Logistic Regression dengan akurasi dan precision masing-masing 77%. Kedua model tersebut terbukti lebih handal dibandingkan Naive Bayes, Random Forest, KNN, maupun Decision Tree. Penelitian ini menegaskan bahwa pendekatan hybrid lexicon-based dan machine learning efektif untuk meningkatkan efisiensi pelabelan sekaligus menghasilkan performa klasifikasi yang optimal pada data ulasan UMKM kuliner, dan diharapkan hasil penelitian ini dapat menjadi referensi bagi pelaku UMKM kuliner dalam memahami persepsi konsumen guna meningkatkan kualitas produk dan layanan secara berkelanjutan.

Kata Kunci: Analisis Sentimen, Lexicon-Based, Machine Learning, UMKM Kuliner, NLP

Abstract

This study aims to classify consumer sentiment in culinary MSMEs in South Kalimantan using a hybrid approach combining lexicon-based methods and machine learning algorithms. The dataset used consists of a collection of consumer reviews from Instagram and TikTok. Lexicon-based is used for automatic labeling based on positive and negative word dictionaries, while six Machine Learning algorithms namely Naive Bayes, Random Forest, Logistic Regression, Support Vector Machine (SVM), Decision Tree, and K-Nearest Neighbors (KNN) are used as pattern-based sentiment classification models. The research stages include dataset retrieval, text preprocessing, automatic labeling, TF-IDF feature extraction, model implementation, and evaluation using accuracy, precision, recall, and F1-score metrics. The results show that SVM is superior with an accuracy of 79% and an F1-score of 70%, followed by Logistic Regression with an accuracy and precision of 77% each. Both models are proven to be more reliable than Naive Bayes, Random Forest, KNN, and Decision Tree. This study confirms that the hybrid lexicon-based and machine learning approach is effective in increasing labeling efficiency while producing optimal classification performance on culinary MSME review data, and it is hoped that the results of this study can be a reference for culinary MSME actors in understanding consumer perceptions to improve product and service quality sustainably.

Keywords: Sentiment Analysis, Lexicon-Based, Machine Learning, Culinary MSMEs, NLP

1. Pendahuluan

Ulasan atau komentar pelanggan yang tersedia di berbagai platform digital—seperti marketplace, media sosial, dan aplikasi layanan makanan—menjadi sumber informasi yang relevan bagi pelaku usaha, termasuk Usaha Mikro, Kecil, dan Menengah (UMKM) di bidang kuliner. Ulasan tersebut mencerminkan tingkat kepuasan pelanggan sekaligus persepsi mereka terhadap kualitas produk dan layanan yang diberikan. Namun, karena volume data yang sangat besar dan berbentuk teks tidak terstruktur, analisis secara manual menjadi tidak praktis dan tidak efisien bagi pelaku UMKM.

Oleh karena itu, dibutuhkan pendekatan berbasis algoritma untuk mengolah dan mengekstrak informasi penting dari data tersebut. Salah satu pendekatan yang umum digunakan adalah klasifikasi sentimen, yaitu metode text mining untuk memetakan ulasan atau opini publik ke dalam label positif, negatif, atau netral (Mardiana et al., 2023). Dalam penerapannya, ketersediaan data berlabel menjadi faktor kunci untuk melatih model machine learning. Proses pelabelan manual membutuhkan sumber daya yang besar, sehingga pendekatan lexicon-based labeling yaitu pelabelan otomatis menggunakan kamus kata bersentimen menjadi solusi yang efisien (Muttakin et al., 2025). Analisis sentimen bertujuan untuk mengklasifikasikan opini ke dalam kategori positif, negatif, atau netral. Dalam penerapannya, ketersediaan data berlabel menjadi tantangan utama karena proses pelabelan manual membutuhkan waktu dan biaya yang besar. Oleh karena itu, digunakan pendekatan lexicon-based labeling yang memungkinkan pelabelan otomatis berdasarkan kamus sentimen (Geni et al., 2023) (Fathoni et al., 2024). Meskipun metode lexicon-based efisien, pendekatan ini memiliki keterbatasan dalam memahami konteks bahasa seperti sarkasme dan ambiguitas. Di sisi lain, algoritma machine learning seperti Naive Bayes, Random Forest, Logistic Regression, Support Vector Machine (SVM), Decision Tree, dan K-Nearest Neighbors (KNN) mampu mempelajari pola data, namun sangat bergantung pada kualitas data latih (Satrya et al., 2023), sering diperlukan analisis perbandingan untuk mengetahui model mana yang paling optimal dalam konteks klasifikasi ulasan UMKM kuliner.

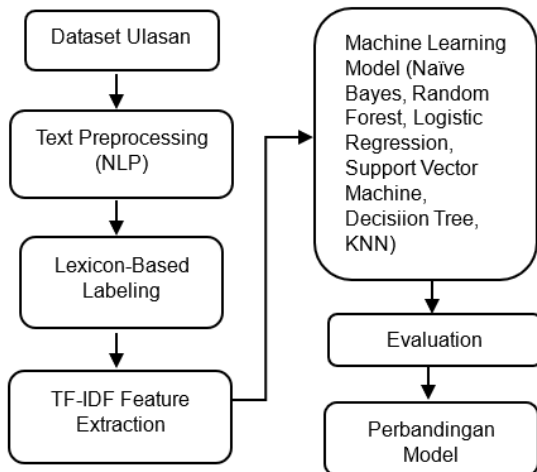
Beberapa penelitian sebelumnya telah mengeksplorasi pendekatan ini. Lexicon-Based digunakan untuk menganalisis 73.120 tweet dan menemukan dominasi opini positif sebesar 77,58%, meski hasilnya sangat bergantung pada

kualitas kamus kata (Hamka & Sari, 2022). Kombinasi Lexicon-based dengan algoritma Support Vector Machine untuk analisis sentimen kebijakan kuliah daring, menghasilkan akurasi 60% dari 400 komentar di platform X. (Oktaviana et al., 2022). Sementara itu, perpaduan Lexicon, Fitur TD-IDF dan Algoritma Naive Bayes pada 741 ulasan restoran dan meraih akurasi 90% (Anggina et al., 2022). Untuk mengetahui kinerja model dalam mengklasifikasikan sentimen terhadap opini publik, peran pelabelan data sangat penting, tidak hanya didasarkan pada nilai atau skor metrik (f1-score) namun juga dipengaruhi oleh distribusi data yang berimbang guna menciptakan model yang objektif dan berlaku umum (Rizkia et al., 2025).

Beberapa penelitian sebelumnya telah menggabungkan lexicon-based dengan machine learning, namun sebagian besar masih terbatas pada satu atau dua algoritma, serta belum memberikan analisis komparatif yang komprehensif antar model pada konteks data ulasan UMKM kuliner. Berdasarkan hal tersebut, research gap dalam penelitian ini adalah belum adanya kajian menyeluruh yang mengintegrasikan lexicon-based labeling dengan perbandingan multi-algoritma machine learning untuk menentukan model terbaik pada klasifikasi sentimen ulasan UMKM kuliner. Evaluasi kinerja klasifikasi dilakukan dengan pendekatan metrik akurasi, precision, recall, dan F1-score untuk memperoleh gambaran kinerja masing-masing model secara menyeluruh (Winata et al., 2023) (Nugraha, 2025). Penelitian ini bertujuan untuk menerapkan lexicon-based labeling sebagai metode pelabelan otomatis serta membandingkan kinerja berbagai algoritma machine learning menggunakan metrik akurasi, precision, recall, dan F1-score. Pendekatan hybrid ini diharapkan mampu menghasilkan kinerja klasifikasi yang optimal, serta memberikan referensi yang bermanfaat bagi pelaku UMKM kuliner di Kalimantan Selatan dalam memahami persepsi konsumen guna meningkatkan kualitas produk dan layanan.

2. Metode Penelitian

Penelitian dilakukan melalui beberapa tahapan sistematis, mulai dari pengambilan dataset hingga evaluasi dan perbandingan kinerja model. Alur penelitian diilustrasikan pada Gambar 1.



Gambar 1. Alur Penelitian

a. Dataset ulasan

Dataset yang digunakan berbentuk file CSV yang diambil dari <https://github.com/setyows/datasetsentimenkuliner>. Dataset ini terdiri dari 3.500 ulasan dalam Bahasa Indonesia dan Bahasa Banjar, berasal dari komentar publik pada platform Instagram dan TikTok milik UMKM kuliner di Kalimantan Selatan, mencakup ulasan positif, negatif, dan netral. Seluruh proses pemrograman dilakukan menggunakan bahasa Python melalui platform Google Colab.

b. Text Preprocessing (NLP)

Dataset ulasan selanjutnya diproses menggunakan teknik Natural Language Processing (NLP) untuk meningkatkan kualitas data sebelum dianalisis. Hal ini bertujuan untuk mengurangi noise dalam data, menyeragamkan format teks serta mengoptimalkan kinerja model klasifikasi (Anggina et al., 2022). Tahapan preprocessing meliputi pembersihan teks (*case folding*, *tokenization*, *stopword removal* dan *stemming*) (Rihastuti & Rosyidi, 2025a).

c. Lexicon-Based Labeling

Data yang telah dibersihkan kemudian dilabeli secara otomatis dengan metode *lexicon-based* (Setiyawan & Mustofa, 2024). Pendekatan ini memanfaatkan kamus sentimen terbotot (positif/negatif) untuk menghitung skor berdasarkan frekuensi kemunculan kata dalam teks. Apabila skor > 0 artinya sentimen positif, jika skor < 0 hasilnya berupa sentimen negatif dan skor = 0 dianggap netral. Kelebihan metode ini adalah prosesnya yang cepat dan tidak membutuhkan anotasi manual. Akan tetapi, kelemahannya terletak pada keterbatasan dalam

memahami konteks bahasa yang kompleks, seperti sarkasme dan makna ambigu, serta ketergantungannya pada kualitas kamus lexicon yang digunakan.

d. TF-IDF Feature Extraction

Data berlabel kemudian diubah menjadi representasi numerik menggunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*) (Rihastuti & Rosyidi, 2025b). TF (*Term Frequency*) mengukur jumlah kemunculan istilah dalam satu dokumen. IDF (*Inverse Document Frequency*) menilai seberapa unik istilah tersebut di seluruh koleksi.

Rumus umum:

$$TF-IDF = TF \times \log \left(\frac{N}{df} \right) \dots \dots \dots (1)$$

Keterangan:

N = jumlah dokumen

df = jumlah dokumen yang mengandung kata tersebut

Metode ini digunakan untuk menonjolkan fitur kata yang relevan sekaligus menekan pengaruh kata-kata yang kurang informatif. Hasilnya berupa matriks fitur numerik yang siap digunakan oleh model machine learning.

e. Penerapan Model Machine Learning

Dataset yang telah direpresentasikan dalam bentuk TF-IDF digunakan untuk melatih enam algoritma klasifikasi, yaitu: (1) Naive Bayes, (2) Random Forest, (3) Logistic Regression, (4) Support Vector Machine (SVM), (5) Decision Tree dan (6) K-Nearest Neighbors (KNN).

Penggunaan beberapa algoritma sekaligus untuk membandingkan kinerja masing-masing model serta menentukan model terbaik untuk kasus analisis sentimen UMKM kuliner. Konsistensi perbandingan dijamin dengan melatih setiap model menggunakan dataset yang serupa.

f. Evaluasi Model

Model yang telah dilatih kemudian dievaluasi menggunakan metrik klasifikasi, yaitu:

- 1) Akurasi yaitu rasio prediksi yang benar
- 2) Precision yaitu ketepatan rasio prediksi positif
- 3) Recall yaitu kemampuan mendeteksi dan menemukan seluruh data positif
- 4) F1-score yaitu metrik keseimbangan antara precision dan recall.

Rumus F1-score:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Evaluasi ini bertujuan untuk mengukur kinerja model secara objektif dan mengetahui kelebihan dan kelemahan masing-masing algoritma.

g. Perbandingan Model

Tahap akhir adalah melakukan analisis perbandingan terhadap seluruh model berdasarkan hasil evaluasi. Tahapan ini meliputi: membandingkan nilai akurasi, precision, recall, dan F1-score, mengidentifikasi model dengan performa terbaik serta menganalisis faktor yang mempengaruhi perbedaan kinerja. Hasil perbandingan ini menjadi dasar dalam penyusunan kesimpulan penelitian.

3. Hasil Penelitian dan Pembahasan

Penelitian ini berfokus pada komentar/ulasan pengguna platform Instagram dan Tiktok dalam pemberian ulasan UMKM kuliner di wilayah Kalimantan Selatan. Setelah dataset diperoleh, kemudian dilakukan text preprocessing, selanjutnya pelabelan otomatis menggunakan pendekatan lexicon-base labeling. Data yang telah diberi label kemudian diekstraksi ke dalam TF-IDF fitur. Berikutnya dilakukan pengujian dengan algoritma machine learning menggunakan Python di Google Colab.

3.1. Penerapan dan Evaluasi Model Machine Learning

a. Naive Bayes

Pengujian menggunakan Naive Bayes menghasilkan output seperti pada Gambar 2.

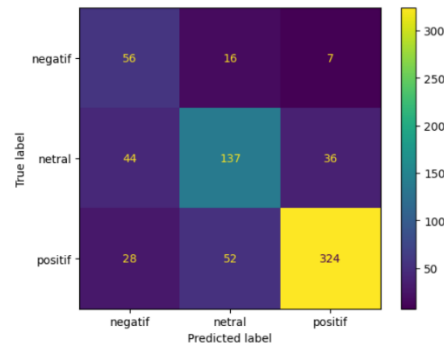
Accuracy: 0.7385714285714285

```
Confusion Matrix
[[ 56 16  7]
 [ 44 137 36]
 [ 28  52 324]]
```

Classification Report				
	precision	recall	f1-score	support
negatif	0.44	0.71	0.54	79
netral	0.67	0.63	0.65	217
positif	0.88	0.80	0.84	404
accuracy			0.74	700
macro avg	0.66	0.71	0.68	700
weighted avg	0.77	0.74	0.75	700

Gambar 2. Akurasi Naive Bayes

Pengujian dengan Naive Bayes menghasilkan akurasi 73,85% (dibulatkan menjadi 74%). Precision 66%, recall 71% dan F1-score 68% (Gambar 2). Dari 700 ulasan pengujian, model berhasil memprediksi 324 ulasan positif dengan benar (true positive), 137 ulasan netral dengan benar 36 diprediksi netral, dan 56 ulasan negatif dengan benar. Meski demikian, masih terdapat 52 ulasan yang sebenarnya positif namun diprediksi netral, 28 ulasan positif yang diprediksi negatif, serta 44 ulasan netral namun diprediksi negatif.



Gambar 3. Confusion Matrix Naive Bayes

Hasil ini sejalan dengan karakteristik Naive Bayes sebagai algoritma generatif yang mengasumsikan independensi antar fitur. Meskipun asumsi tersebut tidak selalu terpenuhi pada data teks nyata, model ini tetap cukup efektif sebagai baseline dan mampu menangkap pola sentimen positif secara umum. Kinerja recall-nya yang relatif tinggi (71%) menunjukkan bahwa model cukup sensitif dalam mendeteksi ulasan positif, meski masih terbatas dalam membedakan sentimen netral dari positif.

b. Random Forest

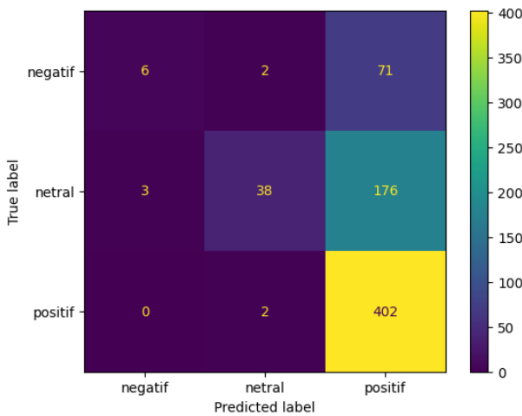
Accuracy: 0.6371428571428571

```
Confusion Matrix
[[ 6  2 71]
 [ 3 38 176]
 [ 0  2 402]]
```

Classification Report				
	precision	recall	f1-score	support
negatif	0.67	0.08	0.14	79
netral	0.90	0.18	0.29	217
positif	0.62	1.00	0.76	404
accuracy			0.64	700
macro avg	0.73	0.42	0.40	700
weighted avg	0.71	0.64	0.55	700

Gambar 4. Akurasi Random Forest

Random Forest menghasilkan akurasi 64%, precision 73%, recall 42% dan F1-score 41% (Gambar 4). Meskipun precision-nya tergolong tinggi, nilai recall yang sangat rendah (42%) menunjukkan bahwa model cenderung hanya melakukan prediksi ketika memiliki keyakinan yang tinggi, akibatnya, banyak ulasan positif dan negatif yang sesungguhnya justru tidak terdeteksi (false negative tinggi). Dari confusion matrix (Gambar 5), terlihat model mampu memprediksi 402 ulasan positif dengan benar, namun salah mengklasifikasikan 176 ulasan netral sebagai positif.



Gambar 5. Confusion Matrix Random Forest

c. Logistic Regression

Model Logistic Regression menghasilkan akurasi sebesar 77%, precision 77%, recall 62% dan F1-score 65% (Gambar 6).

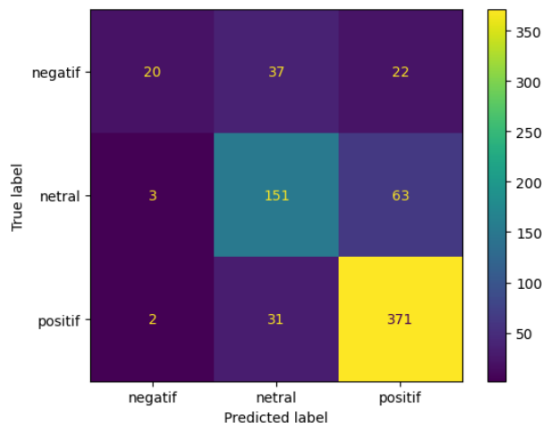
Accuracy: 0.7742857142857142

```
Confusion Matrix
[[ 20  37  22]
 [  3 151  63]
 [  2  31 371]]
```

Classification Report				
	precision	recall	f1-score	support
negatif	0.80	0.25	0.38	79
netral	0.69	0.70	0.69	217
positif	0.81	0.92	0.86	404
accuracy			0.77	700
macro avg	0.77	0.62	0.65	700
weighted avg	0.77	0.77	0.76	700

Gambar 6. Akurasi Logistic Regression

Model ini berhasil memprediksi 371 ulasan positif dengan benar, 151 ulasan netral, dan 20 ulasan negatif. Kesalahan klasifikasi yang terjadi di antaranya 63 ulasan netral yang diprediksi sebagai positif (false positive) dan 31 ulasan positif yang diprediksi sebagai netral (Gambar 7).



Gambar 7. Confusion Matrix Logistic Regression

Logistic Regression dikenal sebagai model linear

yang efisien untuk klasifikasi teks berdimensi tinggi. Ketika dipadukan dengan representasi TF-IDF, model ini mampu memanfaatkan bobot fitur secara langsung untuk membedakan kelas sentimen. Keseimbangan yang baik antara precision dan recall menjadikan Logistic Regression sebagai alternatif yang bagus dan mudah diinterpretasikan, khususnya saat distribusi data antar kelas relatif proporsional.

d. Support Vector Machine

Support Vector Machine menghasilkan akurasi tertinggi sebesar 79%, precision 76%, recall 68% dan F1-score 70% (Gambar 8). Dari 700 ulasan, model berhasil memprediksi 353 ulasan positif dengan benar, 168 ulasan netral, dan 30 ulasan negatif. Kesalahan yang masih terjadi antara lain 43 ulasan netral diprediksi sebagai positif dan 11 ulasan negatif diprediksi sebagai positif (false positive), serta 5 ulasan positif diprediksi sebagai negatif (false negative) seperti terlihat pada Gambar 9.

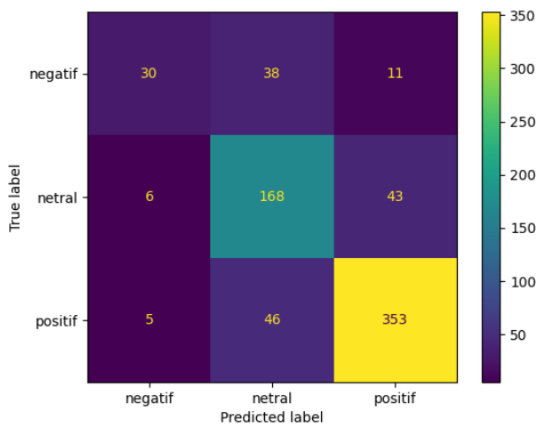
Accuracy: 0.7871428571428571

```
Confusion Matrix
[[ 30  38  11]
 [  6 168  43]
 [  5  46 353]]
```

Classification Report				
	precision	recall	f1-score	support
negatif	0.73	0.38	0.50	79
netral	0.67	0.77	0.72	217
positif	0.87	0.87	0.87	404
accuracy			0.79	700
macro avg	0.76	0.68	0.70	700
weighted avg	0.79	0.79	0.78	700

Gambar 8. Akurasi Support Vector Machine

Keunggulan SVM pada data ini dapat dijelaskan dari karakteristik algoritmanya yaitu SVM bekerja dengan mencari hyperplane optimal yang memaksimalkan margin antar kelas, sehingga sangat efektif pada ruang fitur berdimensi tinggi seperti representasi TF-IDF. Berbeda dengan Naive Bayes yang mengandalkan probabilitas sederhana, SVM mampu menangkap pola yang lebih kompleks dan nonlinear melalui penggunaan kernel. Hal ini sejalan dengan temuan (Oktaviana et al., 2022) dan (Setiyawan & Mustofa, 2024) yang juga melaporkan keunggulan SVM dalam analisis sentimen teks berbahasa Indonesia..



Gambar 9. Confusion Matrix Support Vector Machine

e. Decision Tree

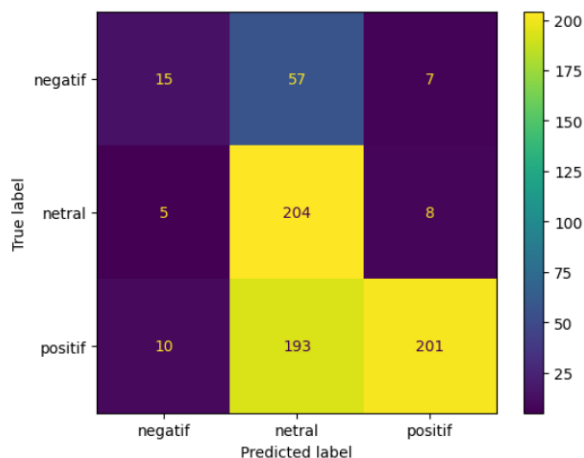
Accuracy: 0.6

Confusion Matrix
[[15 57 7]
[5 204 8]
[10 193 201]]

Classification Report	precision	recall	f1-score	support
negatif	0.50	0.19	0.28	79
netral	0.45	0.94	0.61	217
positif	0.93	0.50	0.65	404
accuracy			0.60	700
macro avg	0.63	0.54	0.51	700
weighted avg	0.73	0.60	0.59	700

Gambar 10. Akurasi Decision Tree

Decision Tree menghasilkan performa terendah di antara semua model, dengan akurasi 60%, precision 63%, recall 54%, dan F1-score 51% (Gambar 10). Meskipun model berhasil memprediksi 201 ulasan positif dan 204 ulasan netral dengan benar, masih terdapat kesalahan yang cukup banyak, seperti 193 ulasan positif yang diprediksi sebagai netral dan berbagai false positive maupun false negative lainnya (Gambar 11).



Gambar 11. Confusion Matrix Decision Tree

Rendahnya kinerja Decision Tree kemungkinan

disebabkan oleh kecenderungan model ini untuk overfitting pada data pelatihan, terutama ketika label yang digunakan—dalam hal ini dihasilkan oleh lexicon-based labeling—mengandung noise. Tanpa mekanisme ensemble, Decision Tree sulit melakukan generalisasi yang baik pada data teks yang kompleks dan tidak terstruktur.

f. K-Nearest Neighbors (KNN)

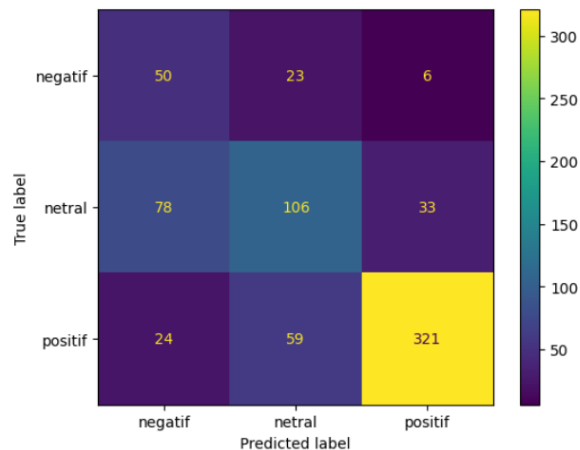
Accuracy: 0.6814285714285714

Confusion Matrix
[[50 23 6]
[78 106 33]
[24 59 321]]

Classification Report	precision	recall	f1-score	support
negatif	0.33	0.63	0.43	79
netral	0.56	0.49	0.52	217
positif	0.89	0.79	0.84	404
accuracy			0.68	700
macro avg	0.59	0.64	0.60	700
weighted avg	0.73	0.68	0.70	700

Gambar 12. Akurasi KNN

KNN menghasilkan akurasi sebesar 68%, precision 59%, recall 64% dan F1-score 60% (Gambar 12). Model berhasil memprediksi dengan benar 321 ulasan positif, 106 ulasan netral, dan 50 ulasan negatif. Namun, terdapat false positive sebanyak 33 ulasan netral yang diprediksi sebagai positif, serta false negative berupa 78 ulasan netral yang diprediksi sebagai negatif (Gambar 13).



Gambar 13. Confusion Matrix KNN

KNN bekerja berdasarkan kedekatan jarak antar sampel, sehingga sensitif terhadap noise dan kurang optimal pada dataset teks berdimensi tinggi. Representasi TF-IDF menghasilkan ruang fitur yang sangat sparse, dan pada kondisi tersebut, pengukuran jarak Euclidean yang digunakan KNN cenderung kehilangan daya diskriminatifnya—fenomena yang dikenal sebagai curse of dimensionality. Hal inilah yang menyebabkan kinerja

KNN berada di posisi pertengahan dibanding model-model lainnya.

3.2. Perbandingan Model

Berdasarkan hasil pengujian seluruh model, berikut adalah tabel perbandingan kinerja algoritma yang disajikan pada Tabel 1.

Tabel 1. Hasil Pengujian menggunakan Machine Learning Model

No	Model	Akurasi	Precision	Recall	F1-score
1	Naive Bayes	74%	66%	71%	68%
2	Random Forest	64%	73%	42%	40%
3	Logistic Regression	77%	77%	62%	65%
4	Support Vector Machine	79%	76%	68%	70%
5	Decision Tree	60%	63%	54%	51%
6	KNN	68%	59%	64%	60%

Berdasarkan Tabel 1, SVM unggul dalam dua metrik utama—akurasi (79%) dan F1-score (70%)—sehingga dinilai sebagai model dengan kinerja terbaik secara keseluruhan. Logistic Regression menyusul dengan akurasi 77% dan precision tertinggi bersama SVM (77%), menunjukkan keseimbangan yang baik antara ketepatan dan sensitivitas. Naive Bayes meraih recall tertinggi (71%), menjadikannya pilihan yang layak sebagai model baseline yang cepat dan sederhana. Di sisi lain, KNN berada di posisi menengah, sementara Random Forest dan Decision Tree mencatat kinerja terendah, terutama akibat nilai recall dan F1-score yang jauh di bawah rata-rata.

Secara keseluruhan, model-model linear seperti SVM dan Logistic Regression terbukti lebih adaptif terhadap noise yang dihasilkan dari proses lexicon-based labeling, dibandingkan model berbasis pohon seperti Decision Tree dan ensemble seperti Random Forest. Temuan ini mengindikasikan bahwa pemilihan algoritma yang tepat perlu mempertimbangkan tidak hanya karakteristik data, tetapi juga kualitas label yang dihasilkan dari metode pelabelan otomatis.

4. Kesimpulan

Berdasarkan hasil penelitian, pendekatan hybrid lexicon-based dan machine learning dinilai efektif untuk meningkatkan efisiensi pelabelan sekaligus menghasilkan performa klasifikasi yang optimal pada data ulasan UMKM kuliner. Pendekatan lexicon-based labeling yang berperan sebagai pelabelan otomatis mempercepat proses anotasi data dan meminimalisir kebutuhan manual labeling. Namun, pendekatan ini berpotensi menghasilkan noise pada label, terutama pada kalimat yang mengandung sarkasme atau konteks ambigu. Dalam pengujian

berbagai algoritma machine learning yang telah dilakukan, model SVM dan Logistic Regression mampu mengatasi keterbatasan tersebut dengan baik. Model linear yaitu Support Vector Machine terbukti sebagai algoritma terbaik dalam mengklasifikasikan sentimen ulasan UMKM kuliner berbasis teks dengan akurasi 79% dan F1-score 70%. Keunggulan SVM terletak pada kemampuannya dalam menentukan batas pemisah antar kelas yang tepat pada ruang fitur berdimensi tinggi, menjadikannya sangat sesuai untuk mengolah data teks yang direpresentasikan dengan TF-IDF. Logistic Regression juga menunjukkan kinerja yang stabil dengan akurasi 77% dan precision 77%. Model berbasis pohon (Decision Tree) dianggap kurang optimal pada dataset hasil lexicon-labeling yang dapat disebabkan overfitting dan ketidakmampuan model dalam menarik kesimpulan secara umum dari data yang kompleks.

5. Daftar Pustaka

- Anggina, S., Setiawan, N. Y., & Bachtiar, F. A. (2022). Analisis Ulasan Pelanggan Menggunakan Multinomial Naïve Bayes Classifier dengan Lexicon-Based dan TF-IDF Pada Formaggio Coffee and Resto. *Ais The Best*, 7, 76–90.
- Fathoni, M. F. N., Puspaningrum, E. Y., & Sihananto, A. N. (2024). Perbandingan Performa Labeling Lexicon InSet dan VADER pada Analisa Sentimen Rohingya di Aplikasi X dengan SVM. *Modem: Jurnal Informatika Dan Sains Teknologi*, 1(3).
- Geni, L., Yulianti, E., & Sensuse, D. I. (2023). Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using IndoBERT Language Models. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI)*, 9(3), 746–757. <https://doi.org/10.26555/jiteki.v9i3.26490>
- Hamka, M., & Sari, D. R. (2022). ANALISIS SENTIMEN DAN INFORMATION EXTRACTION PEMBELAJARAN DARING MENGGUNAKAN PENDEKATAN LEXICON. *Djtechno: Journal of Information Technology Research*, 3(1), 21–32.
- Mardiana, H., Walid, M., & Darmawan, A. K. (2023). LEXICON-BASED AND NAIVE BAYES SENTIMENT ANALYSIS FOR RECOMMENDING THE BEST MARKETPLACE SELECTION AS A MARKETING STRATEGY FOR MSMEs. *PILAR Nusa Mandiri: Journal of Computing and Information System*, 19(2). <https://doi.org/10.33480/pilar.v19i2.4176>
- Muttakin, F., Andrika, N., & Salsabila. (2025). Sentiment Analysis of Shoe Product Reviews on Indonesian E-Commerce Platform Using

- Lexicon Based and Support Vector Machine. *Jurnal Teknik Informatika (Jutif)*, 6(2), 839–854.
- Nugraha, S. A. (2025). PENERAPAN LEXICON BASED UNTUK ANALISIS SENTIMEN MASYARAKAT INDONESIA TERHADAP DANANTARA. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(3), 4949–4957.
- Oktaviana, N. E., Sari, Y. A., & Indriati. (2022). ANALISIS SENTIMEN TERHADAP KEBIJAKAN KULIAH DARING SELAMA PANDEMI MENGGUNAKAN PENDEKATAN LEXICON BASED FEATURES DAN SUPPORT VECTOR MACHINE. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 9(2).
<https://doi.org/10.25126/jtiik.202295625>
- Rihastuti, S., & Rosyidi, A. (2025a). Analisis Sentimen Pengguna Tiktok Tentang Progres Pembangunan Ikn Dengan Metode Random Forest. *Journal of Computer Science and Technology*, 5(1), 19–23.
- Rihastuti, S., & Rosyidi, A. (2025b). Perbandingan Kinerja Support Vector Machine Dan Random Forest Untuk Klasifikasi Sentimen Pengguna Aplikasi Gojek Dengan Optimasi Smote. *Algoritme*, 131–141.
<https://jurnal.mdp.ac.id/index.php/algoritme/article/view/13463>
- Rizkia, A. S., Wufron, & Roji, F. F. (2025). Sentiment Analysis of Coretax : A Comparison of Manual , Transformers- Based , and Lexicon-Based Data Labeling on IndoBERT Performance. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 5(July), 1037–1048.
- Satrya, W. F., Aprilliyani, R., & Yossy, E. H. (2023). Sentiment analysis of Indonesian police chief using multi-level ensemble model ensemble model. *Procedia Computer Science*, 216(2022), 620–629.
<https://doi.org/10.1016/j.procs.2022.12.177>
- Setiyawan, R., & Mustofa, Z. (2024). Comparison of the performance of naive bayes and support vector machine in sirekap sentiment analysis with the lexicon- based approach. *JOSCEX: Journal of Soft Computing Exploration*, 5, 122–132.
- Winata, G. I., Aji, A. F., Cahyawijaya, S., Mahendra, R., Koto, F., Romadhony, A., Kurniawan, K., Moeljadi, D., Prasojo, R. E., Fung, P., Baldwin, T., Lau, J. H., Sennrich, R., & Ruder, S. (2023). NusaX : Multilingual Parallel Sentiment Dataset for 10 Indonesian Local Languages. *Proceedings Ofthe 17th Conference Ofthe European Chapter Ofthe Association for Computational Linguistics*, 815–834.