

## ANALISIS SENTIMEN ULASAN RUMAH MAKAN MENGGUNAKAN PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE* DENGAN *NAÏVE BAYES* (STUDI KASUS: AYAM GORENG NELONGSO CABANG SINGOSARI, MALANG)

Dinar Fairus Salsabillah<sup>1</sup>, Dian Eka Ratnawati<sup>2</sup>, Nanang Yudi Setiawan<sup>3</sup>

<sup>1,2,3</sup>Universitas Brawijaya, Malang

Email: <sup>1</sup>dinarfairus13@student.ub.ac.id, <sup>2</sup>dian\_ilkom@ub.ac.id, <sup>3</sup>nanang@ub.ac.id

\*Penulis Korespondensi

(Naskah masuk: 4 Agustus 2023, diterima untuk diterbitkan: 3 Januari 2024)

### Abstrak

Peningkatan kualitas produk dan pelayanan merupakan tantangan yang dihadapi oleh bisnis kuliner, termasuk rumah makan Ayam Goreng Nelongso Singosari di Kabupaten Malang. Analisis sentimen digunakan untuk mengidentifikasi ulasan pelanggan terkait pelayanan, kualitas produk, harga, dan kepuasan pelanggan. Penelitian ini membandingkan metode *Support Vector Machine* (SVM), metode *Naïve Bayes Classifier* (NBC), dan *Root Cause Analysis* untuk mengklasifikasikan sentimen ulasan pelanggan dan menganalisis masalah yang mendasarinya. Tujuan penelitian ini adalah membandingkan hasil dari algoritma *Support Vector Machine* dan *Naïve Bayes Classifier* dalam pengklasifikasian sentimen ulasan pelanggan rumah makan Ayam Goreng Nelongso. Penelitian ini juga bertujuan untuk menghasilkan rekomendasi berdasarkan analisis *root cause* pada sentimen negatif ulasan pelanggan. Implementasi kedua algoritma klasifikasi menunjukkan performa yang baik dalam mengklasifikasikan data dengan akurasi tinggi. Pengujian menunjukkan bahwa kinerja SVM lebih unggul dengan tingkat akurasi mencapai 92,74%, sementara NBC mencapai tingkat akurasi sebesar 91,67%. Hasil analisis *root cause* menunjukkan beberapa rekomendasi untuk meningkatkan aspek harga, makanan, layanan, dan tempat rumah makan. Rekomendasi yang dapat dilakukan oleh pihak rumah makan diantaranya adalah evaluasi ukuran dan harga, penggunaan *deep frying*, pelatihan dan evaluasi pelayanan, serta penambahan tenaga kerja atau kerjasama dengan outsourcing dalam menjaga kebersihan tempat. Hasil penelitian diharapkan dapat membantu pemilik rumah makan dalam mengembangkan kualitas produk dan pelayanan serta memberikan pandangan untuk langkah-langkah yang dapat diambil di kemudian hari.

**Kata kunci:** Rumah Makan, Analisis Sentimen, *Naïve Bayes*, *Support Vector Machine*, *Root Cause Analysis*

## SENTIMENT ANALYSIS OF RESTAURANT REVIEWS USING COMPARISON OF *SUPPORT VECTOR MACHINE* ALGORITHM WITH *NAÏVE BAYES* (CASE STUDY: AYAM GORENG NELONGSO SINGOSARI BRANCH, MALANG)

### Abstract

The improvement of product quality and services poses a challenge for culinary businesses, including Nelongso Singosari Fried Chicken restaurant in Malang Regency. Sentiment analysis is employed to identify customer reviews related to service, product quality, pricing, and customer satisfaction. This research compares the *Support Vector Machine* (SVM) and *Naïve Bayes Classifier* (NBC) methods, along with *Root Cause Analysis*, to classify customer sentiment and analyze underlying issues. The aim is to compare the results of the *Support Vector Machine* and *Naïve Bayes Classifier* algorithms in classifying customer sentiment for Nelongso Fried Chicken restaurant. The research also seeks to generate recommendations based on root cause analysis of negative customer reviews. The implementation of both classification algorithms demonstrates good performance in classifying data with high accuracy. Testing indicates that the SVM outperforms with an accuracy rate of 92.74%, while the NBC achieves an accuracy rate of 91.67%. Root cause analysis results provide several recommendations to enhance pricing, food, service, and the restaurant's ambiance. Recommendations for the restaurant include evaluating portion sizes and pricing, optimizing deep frying techniques, conducting training and service evaluations, and considering additional staff or outsourcing for cleanliness maintenance. The research findings are expected to assist restaurant owners in improving product and service quality and offer insights for future actions.

**Keywords:** *Restaurant, Sentiment Analysis, Naïve Bayes, Support Vector Machine, Root Cause Analysis*

## 1. PENDAHULUAN

Usaha Kuliner memikat banyak orang sebagai salah satu bidang usaha yang menarik, terbukti dari banyaknya restoran yang menawarkan berbagai produk baik dari segi harga maupun pelayanan. Pemilik rumah makan perlu untuk melakukan inovasi guna meningkatkan mutu produk sehingga pelanggan merasa puas dengan produk yang dibeli. Seiring dengan ketatnya persaingan di bisnis kuliner, adaptasi produk dengan selera pelanggan menjadi kunci untuk tetap bersaing (Yazid, 2008). Pelanggan tidak hanya mencari produk sesuai selera, tetapi juga pelayanan yang memuaskan. Kualitas pelayanan yang baik, termasuk tempat yang nyaman dan bersih serta tersedianya fasilitas yang memenuhi kebutuhan pelanggan, berdampak positif pada kepuasan pelanggan (Maulidah, Widodo and Zulianto, 2019).

Rumah makan Ayam Goreng Nelongso Singosari adalah salah satu bisnis kuliner yang berada di Kab.Malang, Jawa Timur. Rumah makan ini memiliki produk unggulan yaitu ayam yang dijual dengan berbagai variasi sambal yang tidak dapat ditemukan di rumah makan lain, hal ini tentu menjadi daya tarik bagi pelanggan. Rumah Makan ini juga menawarkan pilihan menu lain seperti ikan, bebek, berbagai sayuran, dan minuman yang dapat dinikmati oleh para pelanggan. Beberapa fasilitas juga disediakan, seperti wifi, musholla, tempat parkir, hingga pelayanan selama 24 jam. Banyaknya bisnis kuliner terutama yang menjual produk olahan ayam di kota Malang, menyebabkan persaingan di industri ini semakin ketat. Ayam Goreng Nelongso harus mempertahankan kualitas dan pelayanan produk yang diberikan kepada pelanggan agar dapat bertahan dalam bisnis ini.

Permasalahan yang dihadapi oleh pemilik Rumah Makan Ayam Goreng Nelongso Singosari adalah kesulitan untuk mengetahui opini atau ulasan pelanggan secara efisien. Jumlah ulasan yang banyak membuat sulit bagi pemilik usaha untuk secara manual memahami dan menganalisis setiap ulasan pelanggan. Selain itu, pemilik usaha juga menghadapi tantangan dalam mengidentifikasi aspek-aspek yang harus dibenahi berdasarkan ulasan pelanggan tersebut. Salah satu cara untuk mengidentifikasi pelayanan dan kualitas produk yang diinginkan pelanggan dapat dilihat dari ulasan pelanggan Ayam Goreng Nelongso cabang Singosari yang pernah berkunjung yakni dengan menggunakan analisis sentimen.

Analisis sentimen adalah proses menganalisis dan mengekstraksi opini dari teks atau data yang tidak terstruktur secara otomatis, guna mendapatkan informasi mengenai sentimen (Kaura et al., 2015). Analisis sentimen dapat menghasilkan pola dari teks yang diperoleh dari dokumen, sehingga memungkinkan untuk menentukan apakah pendapat

dari dokumen masuk kategori positif atau negatif (Rofiqoh, Setya Perdana and Fauzi, 2017).

Metode *Machine Learning* digunakan untuk mengidentifikasi review pelanggan terhadap Ayam Goreng Nelongso Singosari. Algoritma NBC dan SVM digunakan untuk menganalisis dan mengevaluasi review pelanggan berdasarkan data yang terkumpul. Sebelumnya, Nanang telah menjalankan studi mengenai evaluasi sentimen yang terkait dengan penerapan sistem pelat nomor ganjil dan genap di platform Twitter menggunakan metode NBC, dan berhasil mencapai tingkat ketepatan sebesar 86,67% (Ruhayana, 2019). Penelitian lain oleh Sifa Amalia, dkk membahas jasa pesan antar makanan pada restoran Solaria menggunakan metode SVM dan KNN, dan hasil pengujian menunjukkan bahwa SVM mempunyai hasil yang lebih unggul dengan akurasi sebesar 81,92% (Sifa Amalia et al., 2021). Penelitian lain mengenai ulasan beberapa rumah makan oleh Yoel et al., menggunakan algoritma *Support Vector Machine* pada ulasan pelanggan di *web*, dengan hasil akurasi dan *f-measure* yang tinggi sebesar 93% (Julianto, Setiabudi and Rostianingsih, 2022).

Algoritma *Naïve Bayes* akan dibandingkan dengan *Support Vector Machine* untuk memilih algoritma yang memiliki hasil klasifikasi terbaik. Pemilihan metode-metode ini didasarkan pada karakteristik masing-masing. *Naïve Bayes* dipilih karena sederhana dan memiliki waktu pemrosesan yang cepat, sementara *Support Vector Machine* dipilih untuk mengatasi data yang kompleks dan hubungan non-linear. Tujuan membandingkan kedua metode ini adalah mengetahui dan menentukan algoritma yang memberikan hasil klasifikasi terbaik untuk ulasan pelanggan di sektor bisnis kuliner. Penelitian ini berfokus pada sektor bisnis kuliner, khususnya Rumah Makan Ayam Goreng Nelongso Singosari di Kabupaten Malang, Jawa Timur. Penelitian sebelumnya mungkin telah mengeksplorasi analisis sentimen pada industri yang berbeda atau tidak memiliki fokus pada bisnis kuliner spesifik.

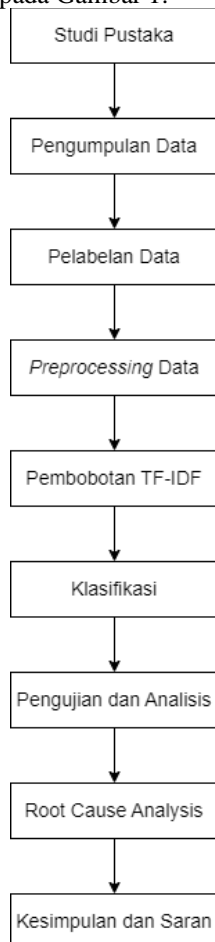
Kontribusi penelitian ini terletak pada pengembangan metode analisis sentimen untuk meningkatkan pemahaman terhadap preferensi pelanggan terhadap pelayanan, produk, harga, dan kepuasan di Rumah Makan Ayam Goreng Nelongso Singosari. Melalui pemilihan algoritma terbaik, diharapkan penelitian ini memberikan wawasan baru dan dapat membantu pemilik rumah makan dalam mengidentifikasi opini pelanggan dengan lebih efisien, meningkatkan kualitas pelayanan dan produk, serta memperbaiki kepuasan pelanggan secara keseluruhan. Penelitian ini tidak hanya memberikan solusi praktis untuk Rumah Makan Ayam Goreng Nelongso Singosari dalam mengelola ulasan

pelanggan, tetapi juga memberikan kontribusi pada pengembangan pengetahuan terkait analisis sentimen di industri kuliner. Melalui fokus pada keunggulan metode, keterkaitan dengan praktik bisnis, dan implikasi praktis, penelitian ini diharapkan dapat menjadi panduan berharga untuk pemilik usaha kuliner lainnya dalam meningkatkan kualitas layanan dan produk mereka.

Berdasarkan penjelasan pada latar belakang, diusulkan sebuah penelitian analisis sentimen terkait aspek pelayanan, produk, harga, serta kepuasan pelanggan pada objek Ayam Goreng Nelongso cabang Singosari Malang. Penelitian ini menggunakan algoritma *Support Vector Machine*, *Naïve Bayes Classifier*, dan *Root Cause Analysis* (RCA) sebagai metode analisis masalah. Diharapkan dengan adanya penelitian ini mempermudah pemilik rumah makan untuk mengembangkan kualitas produk dan pelayanan serta dapat mengetahui langkah apa yang dapat dilakukan kedepannya.

## 2. METODE PENELITIAN

Tahapan atau kerangka berpikir pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

### 2.1. Pengumpulan Data

Data ulasan pada rumah makan Ayam Goreng Nelongso dikumpulkan menggunakan teknik *web scraping* dengan bantuan *extension Google Chrome* yaitu *instan data scraper*. Data yang diambil sebanyak 741 ulasan dari tiga tahun lalu terhitung dari Bulan Mei 2020 sampai Bulan Mei 2023.

### 2.2. Pelabelan Data

Pelabelan sentimen dan aspek akan dilakukan secara manual. Label sentimen memiliki tiga kategori yaitu positif, negatif, dan positif negatif, sedangkan label aspek memiliki empat kategori yaitu makanan, layanan, harga, dan tempat. Data pada proses pelabelan melalui beberapa tahap yaitu memisah ulasan yang mengandung positif negatif dari yang jumlahnya 741 ulasan menjadi 824 ulasan dengan kategori positif dan negatif saja. Hasil dari pelabelan sentimen yang berjumlah 824 ulasan akan dilabeli aspek dengan jumlah akhir yaitu 886 ulasan.

### 2.3. Text Mining

*Text mining* merupakan proses pengelolaan data dengan tujuan untuk mengidentifikasi dan memberikan arti pada data yang tidak teratur atau terstruktur (Jo, 2019).

### 2.4. Preprocessing

*Preprocessing* data bertujuan untuk membersihkan dan mempersiapkan data untuk diolah pada proses pembobotan. Pada tahap ini akan dilakukan lima proses dengan proses awal adalah *case folding*, mengubah huruf kapital menjadi huruf kecil dan menghapus karakter merupakan langkah pertama. Langkah kedua adalah tokenisasi, yaitu mengubah kalimat menjadi unit yang lebih kecil. Ketiga, *normalizing* yaitu mengubah kata tidak baku menjadi baku. Keempat, *stemming* yaitu menghapus imbuhan awal dan akhir. Kelima, *filtering* yaitu menghapus kata yang tidak diperlukan dengan mengecualikan beberapa kata dalam *stopword learning* (Raschka et al., 2020).

### 2.5. Pembobotan TF-IDF

Klasifikasi tidak dapat memproses langsung dokumen dalam bentuk aslinya, sehingga diperlukan pembobotan yang bertujuan untuk mencari bobot kata yang muncul dalam sekumpulan dokumen (Jo, 2019). Pembobotan TF-IDF dapat dihitung dan ditentukan dengan menggunakan Persamaan 1.

$$TF\text{-}IDF = tf_{t,d} \times \log\left(\frac{N}{df_{t,N}}\right) \quad (1)$$

Keterangan:

$tf_{t,d}$  = Frekuensi *term* (kata) dalam dokumen

$N$  = Jumlah dokumen

$df_{t,N}$  = Jumlah dokumen yang mengandung kata

### 2.6. Naïve Bayes Classifier

Metode klasifikasi probabilistik sederhana yang disebut sebagai *Naïve Bayes* memiliki kecepatan pemrosesan yang tinggi dan mengasumsikan independensi antar fitur dalam data (Kaushik et al., 2022). Penelitian ini akan menggunakan jenis *Naïve Bayes* yang disebut *Multinomial Naïve Bayes*. Perhitungan *Naïve Bayes* dapat dilakukan dengan menghitung probabilitas kategori positif dan negatif (*prior*) menggunakan Persamaan 2, kemudian menghitung atribut (*likelihood*) menggunakan Persamaan 3, dan menghitung probabilitas dokumen dengan atribut masuk kedalam kategori positif atau negatif menggunakan Persamaan 4.

$$\text{Prior} = \frac{\text{total(positif/negatif)}}{\text{total kelas}} \quad (2)$$

$$P(W_i|C_j) = \frac{W_{cj}+1}{\sum w' \epsilon v W_{cj}'+B'} \quad (3)$$

$$\text{Posterior} = \text{Prior} \times \text{Likelihood} \quad (4)$$

### 2.7. Support Vector Machine

Metode pengklasifikasian yang efisien dan cepat dalam menetapkan label pada teks dengan pendekatan geometris biasa disebut *Support Vector Machine*. Cara kerja *Support Vector Machine* adalah mencari *hyperplane* yang paling optimal untuk memaksimalkan margin. *Hyperplane* berfungsi untuk memisahkan kategori kelas positif dan kelas negatif (Jo, 2019). Dalam metode SVM juga terdapat fungsi kernel  $K(x_i, x)$  yang berfungsi untuk mentransformasikan data ke ruang berdimensi tinggi (Simangunsong et al., 2019). Tipe kernel yang digunakan pada penelitian ini adalah *Radial Basis Function* (RBF). Langkah-langkah pelatihan dalam SVM menggunakan *Sequential Learning* (Zaiem Praghakusma & Charibaldi, 2021) yaitu:

1. Inisialisasi parameter:  $\alpha_i, \gamma, \epsilon$ , dan C,

Keterangan:

$\alpha$  = nilai alfa

$\gamma$  = nilai gamma

C = *complexity*

E = nilai epsilon

2. Menghitung perhitungan kernel

$$K(x_i, x) = \exp(-\gamma |x - x_i|^2 + C) \quad (5)$$

3. Menghitung matriks *Hessian*

$$D_{ij} = y_i y_j (K(x_i, x_j) + \lambda^2) \quad (6)$$

Keterangan:

$x_i$  = data ke-i

$x_j$  = data ke-j

$y_i$  = kelas data ke-i

$y_j$  = kelas data ke-j

$\lambda$  = lambda

$K(x_i, x_j)$  = fungsi kernel.

4. Menghitung nilai error sebagai berikut:

$$E_i = \sum_{j=1}^n \alpha_j [D]_{ij} \quad (7)$$

Keterangan:

$\alpha_j$  = alfa ke-j

$[D]_{ij}$  = matriks hessian

$E_i$  = error rate

5. Menghitung  $\delta\alpha_i$

$$\delta\alpha_i = \min\{\max[\gamma(1 - E_i), -\alpha_i], C - \alpha_i\} \quad (8)$$

Keterangan:

$\delta\alpha_i$  = alfa ke-i

$\gamma$  = gamma

$E_i$  = error rate

C = *complexity*

6. Menghitung  $\alpha_i$  baru

$$\alpha_i = \delta\alpha_i + \alpha_i \quad (9)$$

Keterangan:

$\alpha_i$  = alfa ke-i

$\delta\alpha_i$  = delta alfa ke-i

7. Menghitung nilai bias

$$b = -\frac{1}{2} (\sum_{i=1}^n \alpha_i y_i K(x_i, x^-) + \sum_{i=1}^n \alpha_i y_i K(x_i, x^+)) \quad (10)$$

Keterangan:

$\alpha_i$  = alfa ke-i

$y_i$  = kelas data ke-i

$K(x_i, x^-)$  = fungsi kernel data negatif.

$K(x_i, x^+)$  = fungsi kernel data positif.

8. Menghitung fungsi  $f(x)$ :

$$f(x) = \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \quad (11)$$

Keterangan:

b = parameter *hyperplane* yang dicari

$\alpha_i$  = alfa ke-i

$y_i$  = kelas data ke-i

$K(x_i, x_j)$  = fungsi kernel.

### 2.8. Confusion Matrix

*Confusion matrix* adalah metode evaluasi yang bertujuan untuk menguji efektivitas model *machine learning* dengan memeriksa dan membandingkan prediksi model dengan nilai yang sebenarnya. *Confusion matrix* dapat membantu dalam memahami jenis kesalahan yang dilakukan oleh model, seperti

*false positive* dan *false negative* (Raschka et al., 2020). *Confusion matrix* adalah tabel yang menampilkan empat *matrix* evaluasi yaitu *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN).

Tabel 1. *Confusion Matrix*

	<i>Predicted Negatives</i>	<i>Predicted Positives</i>
<i>Actual Negative</i>	<i>True Negative (TN)</i>	<i>False Negative (FN)</i>
<i>Actual Positives</i>	<i>False Positif (FP)</i>	<i>True Positive (TP)</i>

- Accuracy adalah nilai dari kelas prediksi benar dibagi dengan seluruh nilai yang ada, accuracy menggambarkan kinerja model dalam melakukan klasifikasi dengan benar.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (12)$$

- Precision adalah nilai prediksi benar pada kelas positif dibagi dengan total nilai prediksi pada kelas positif. Precision menggambarkan kemampuan sistem dalam mencari ketepatan informasi yang diminta.

$$Precision = \frac{TP}{TP+FP} \quad (13)$$

- *Recall* dapat dijelaskan sebagai rasio antara prediksi yang benar terhadap kelas positif dengan jumlah prediksi yang benar terhadap kelas positif dan kesalahan prediksi terhadap kelas negatif. *Recall* mencerminkan kemampuan sistem dalam mendeteksi kembali suatu informasi.

$$Recall = \frac{TP}{TP+FN} \quad (14)$$

- *F-measure* adalah nilai rata-rata antara presisi dan recall, semakin mendekati angka 1 maka semakin optimal hasilnya.

$$F\text{-measure} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (15)$$

## 2.9. K-Fold Cross Validation

*K-fold cross-validation* merupakan metode evaluasi yang digunakan untuk menguji kinerja model pada sejumlah sampel data yang terbatas. Teknik ini melibatkan pembagian data menjadi k subset yang memiliki ukuran yang sama, atau disebut "*fold*", dan melatih model sebanyak k kali. Setiap kali melatih model, salah satu *fold* digunakan sebagai data validasi, sementara *fold* yang lain digunakan sebagai data latihan. Performa model kemudian dievaluasi dengan cara menghitung rata-rata hasil dari k

pelatihan yang dilakukan (Chamorro-Atalaya et al., 2023).

## 2.10. Root Cause Analysis

*Root Cause Analysis* (RCA) merupakan metode untuk menemukan dan mengenali sumber masalah secara mendalam dengan tujuan mencegah terjadinya masalah di kemudian hari (Park and Kim, 2021). Implementasi RCA melibatkan *data collection*, *causal factor chart*, *root cause identification*, dan *recommendation generation*. Tahap *data collection* dilakukan untuk mengumpulkan bukti tentang masalah yang terjadi, tahap *causal factor chart* melibatkan penggunaan *word cloud*, tahap *root cause identification* melibatkan metode *5 whys* untuk mengidentifikasi akar masalah. Terakhir, *recommendation generation* diterapkan untuk mencegah terjadinya masalah serupa di kemudian hari (C., 2021).

## 3. HASIL DAN PEMBAHASAN

### 3.1. Pelabelan Data

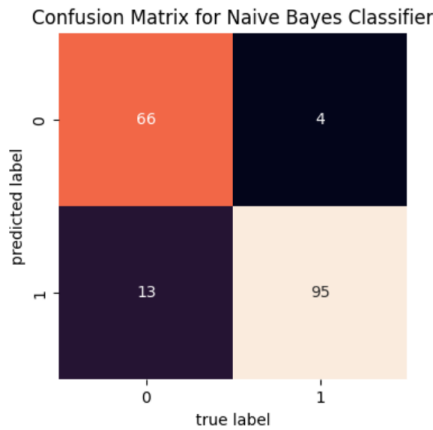
Hasil dari pelabelan data menghasilkan total data akhir sebanyak 886 ulasan dengan distribusi data setiap aspek nya dapat dilihat pada Tabel 3.

Tabel 2. *Recommendation Generation*

Aspek	Positif	Negatif	Total
Makanan	197	237	434
Layanan	71	83	154
Harga	114	11	125
Tempat	123	50	173
Jumlah	505	381	886

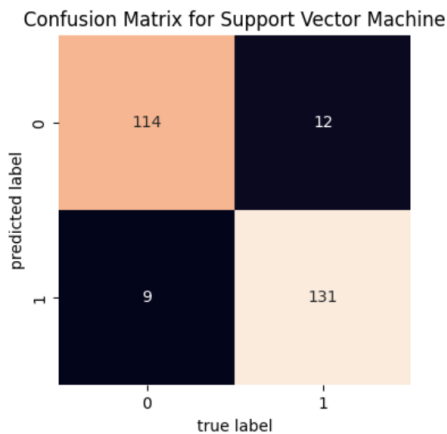
### 3.2. Confusion Matrix

Pengujian performa algoritma dilakukan untuk mengevaluasi model. Pengujian menggunakan *confusion matrix 2x2* yang terdiri dari kelas prediksi dan kelas aktual. Pada pengujian *Naïve Bayes*, data pelatihan dan data pengujian dipisahkan dengan perbandingan 80%:20%, yang mana terdapat 708 ulasan dalam data pelatihan dan 178 ulasan dalam data pengujian. Pada pengujian metode *Support Vector Machine* untuk pengujian, pembagian data pelatihan dan pengujian disesuaikan menjadi 70%:30%, dengan jumlah data pelatihan mencapai 620 ulasan dan data pengujian mencapai 266 ulasan. Hasil pengujian *Confusion Matrix* pada *Naïve Bayes* dan SVM dapat dilihat pada Gambar 2 dan 3.



Gambar 2. Confusion Matrix Naive Bayes

Dari pengujian 79 ulasan negatif, sistem mengklasifikasikan ulasan dengan *true negative* sebanyak 66 dan *false negative* sebanyak 13, sedangkan dari 99 ulasan positif, sistem mengklasifikasi ulasan dengan *true positive* sebanyak 95 dan *false positive* sebanyak 4.

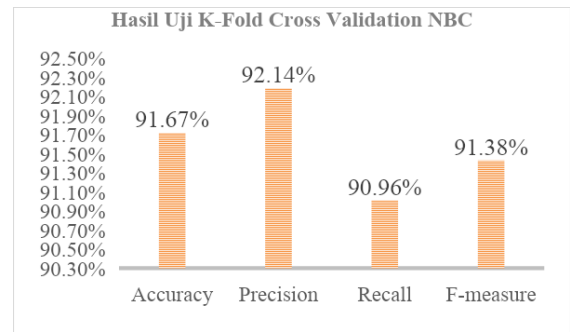


Gambar 3. Confusion Matrix Support Vector Machine

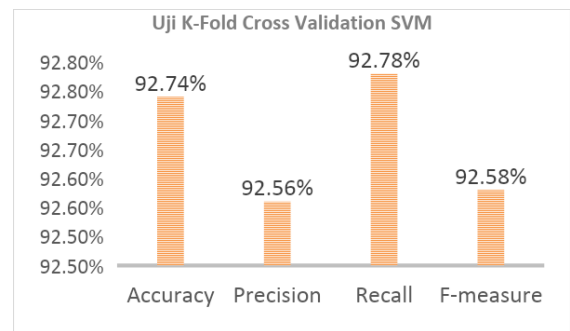
Dari pengujian 123 ulasan negatif, sistem mengklasifikasikan ulasan dengan *true negative* sebanyak 114 dan *false negative* sebanyak 9, sedangkan dari 143 ulasan positif, sistem mengklasifikasi ulasan dengan *true positive* sebanyak 131 dan *false positive* sebanyak 12.

### 3.3. K-Fold Cross Validation

*K-fold cross validation* diterapkan untuk memvalidasi data dan meningkatkan efektivitas pengujian. Penggunaan nilai *k* sebesar 10 mengindikasikan bahwa pengujian dilakukan dalam 10 iterasi berbeda. Rata-rata hasil dari setiap iterasi dijadikan sebagai output akhir pengujian. Hasil evaluasi menggunakan teknik *k-fold cross validation* pada metode NBC dan SVM dapat dilihat dalam Gambar 4 dan Gambar 5.



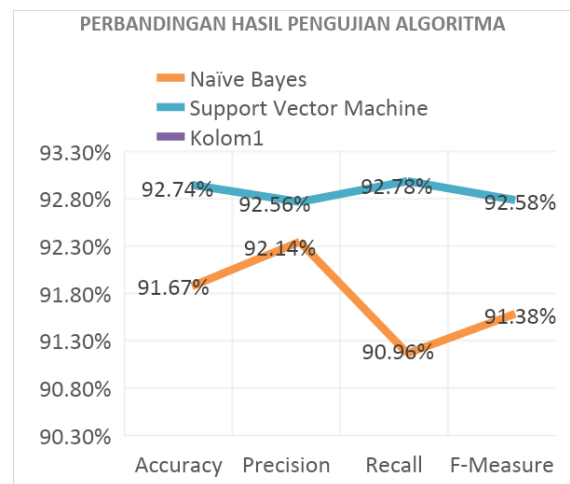
Gambar 4. Uji K-Fold Naive Bayes



Gambar 5. Uji K-Fold Support Vector Machine

### 3.4. Perbandingan Hasil NBC dan SVM

Perbandingan dilakukan untuk mengetahui algoritma klasifikasi terbaik berdasarkan klasifikasi yang telah dilakukan. Pengujian dilakukan dengan menggunakan rasio data pelatihan dan pengujian sebesar 80%:20% untuk *Naive Bayes* dan 70%:30% untuk *Support Vector Machine*, serta menggunakan pengujian sebanyak 10 iterasi untuk masing-masing algoritma. Perbandingan hasil mencakup nilai akurasi, presisi, *recall*, dan *f-measure* yang ditampilkan pada Gambar 6.



Gambar 6. Perbandingan Pengujian NBC dan SVM

### 3.5. Root Cause Analysis

#### a. Data Collection

Data yang digunakan pada analisis *root cause* adalah data sentimen negatif pada setiap aspek dengan rincian pada Tabel 3.

#### b. Causal Factor Chart

Data yang telah telah dikumpulkan selanjutnya akan dianalisis untuk mengidentifikasi pola yang mungkin berkaitan dengan masalah. Proses *causal factor* akan menggunakan bantuan *word cloud* untuk melihat kata atau faktor yang paling sering muncul sehingga akan mendapatkan perhatian penuh untuk diteliti. *Word cloud* pada setiap aspek ditampilkan pada Gambar 7, 8, 9, dan 10.



Gambar 7. Word Cloud Aspek Harga



Gambar 8. Word Cloud Aspek Layanan



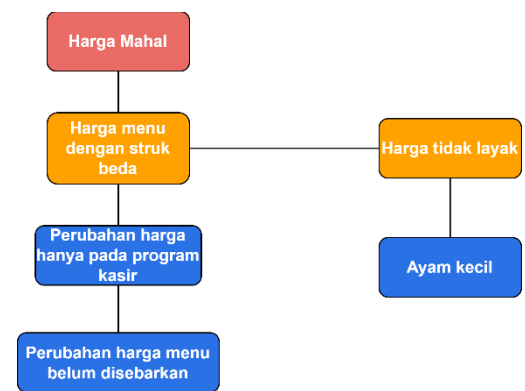
Gambar 9. Word Cloud Aspek Makanan



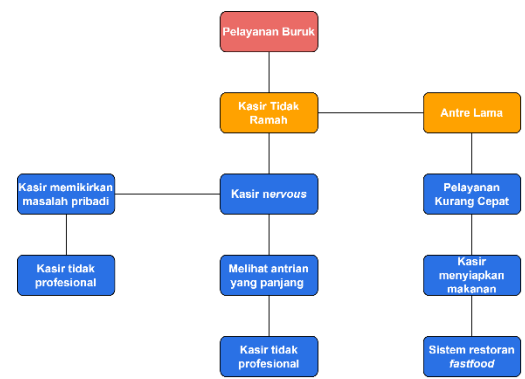
Gambar 10. Word Cloud Aspek Tempat

#### c. Root Cause Identification

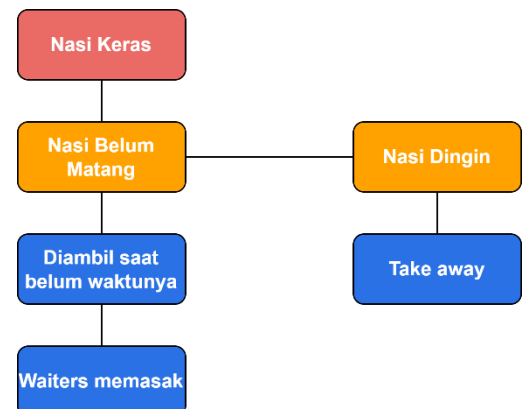
Proses menemukan dan menganalisis masalah akan menggunakan metode *focus group discussion* dengan pihak rumah makan untuk menemukan akar permasalahan serta memvalidasi data hasil *causal factor chart*. Tahap ini akan menganalisis kata yang sering muncul pada *word cloud* pada proses sebelumnya dan menggunakan metode *5 whys* untuk mengidentifikasi faktor penyebab dengan menanyakan “mengapa?” sampai ke akar masalahnya. Gambar 11, 12, 13, 14, dan 15 adalah hasil metode *5 whys*.



Gambar 11. 5 Whys Root Cause Aspek Harga

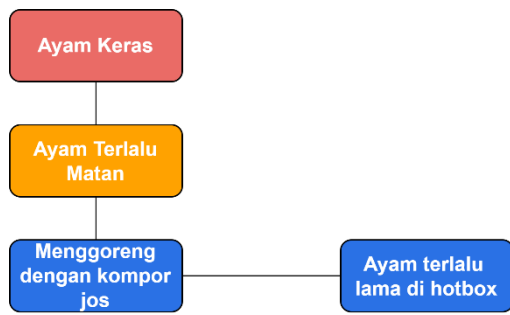


Gambar 12. 5 Whys Root Cause Aspek Layanan

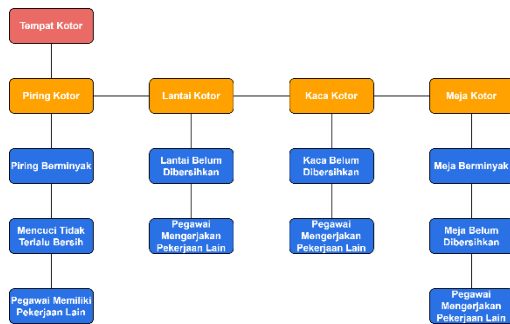


Gambar 13. 5 Whys Root Cause Aspek Makanan (1)





Gambar 14. 5 Whys Root Cause Aspek Makanan (2)



Gambar 15. 5 Whys Root Cause Aspek Tempat

d. Recommendation Generation

Recommendation generation merupakan proses identifikasi solusi yang dapat diambil untuk menangani akar masalah yang dihasilkan dari proses sebelumnya. Tahap untuk menemukan solusi dan rekomendasi yang diberikan melibatkan pihak rumah makan yaitu head manager dan trainee manager untuk menghasilkan solusi efektif dan meminimalisir kesalahan dikemudian hari. Rekomendasi yang diberikan berdasarkan root cause yang ada ditampilkan pada Tabel 3.

Tabel 3. Recommendation Generation

Aspek	Root Cause	Rekomendasi
Harga	Perubahan harga menu belum disebarkan	Menyebarkan informasi perubahan segera setelah terdapat perubahan harga.
	Ayam kecil	Evaluasi ukuran dan harga menu yang ada.
Layanan	Kasir tidak profesional	Melakukan pelatihan pegawai untuk mengatasi kecemasan dan gugup saat bekerja.
	Sistem rumah makan fast food	Menyempurnakan sistem fastfood. Evaluasi kebutuhan tenaga kerja pada jam-jam sibuk seperti sore hari dan hari weekend. Manajemen ekspektasi pelanggan
Makanan	Waiters memasak	Penempatan tugas yang sesuai. Pengawasan dan pemantauan.
	Sistem take away	Evaluasi sistem take away.

	Penggunaan kompor jos	Pertimbangkan untuk mengubah metode kepada metode yang lebih terkontrol misalnya deep frying.
Tempat	Ayam terlalu lama di hotbox	Kelola suhu dan waktu di hotbox.
	Pegawai memiliki pekerjaan lain	Penambahan petugas kebersihan dan melakukan audit kebersihan.

4. PENUTUP

4.1. Kesimpulan

Berdasarkan pengujian yang telah dilakukan, didapatkan kesimpulan sebagai berikut:

1. Implementasi dua algoritma klasifikasi menunjukkan metode NBC dan SVM memiliki performa yang bagus dalam mengklasifikasikan data dengan benar, hal ini dibuktikan dengan hasil akurasi keduanya yang tinggi. Hasil pengujian dan evaluasi menggunakan confusion matrix dan k-fold cross validation menunjukkan bahwa SVM memberikan hasil yang lebih unggul apabila dibandingkan dengan algoritma Naïve Bayes Classifier. Hasil dari pengujian Naïve Bayes diperoleh nilai akurasi sebesar 91.67%, presisi 92.14%, recall 90.96%, dan f-measure 91.38%, sedangkan Support Vector Machine mendapatkan hasil yang lebih unggul dengan nilai accuracy 92.74%, precision 92.56%, recall 92.78%, dan f-measure 92.58%.
2. Berdasarkan wawancara dan diskusi dengan manajer cabang Ayam Goreng Nelongso Singosari, recommendation generation yang dapat dilakukan untuk aspek harga adalah evaluasi ukuran dan harga, pihak rumah makan berencana untuk menambah ukuran ayam. Pada aspek makanan rekomendasi yang dapat dilakukan adalah mengganti penggunaan kompor jos menjadi deep frying, selain itu rekomendasi penempatan kerja yang sesuai juga sudah dilakukan. Pada aspek layanan, rekomendasi yang dapat dilakukan adalah meningkatkan pelatihan dan evaluasi setiap bulannya agar menjaga kualitas pelayanan dari rumah makan. Pada aspek tempat rekomendasi yang dapat dilakukan adalah menambah tenaga kerja baru atau bekerja sama dengan outsourcing dalam segi kebersihan tempat.

4.2. Prospek Penelitian Lanjutan

Berikut adalah beberapa prospek penelitian lanjutan yang dapat penulis berikan berdasarkan temuan dalam penelitian ini:



1. Melakukan klasifikasi pada level aspek sehingga penelitian lebih terfokus pada aspek yang diteliti.
2. Membuat visualisasi *dashboard* untuk mempermudah penyampaian hasil klasifikasi kepada pihak rumah makan.

#### DAFTAR PUSTAKA

- Chamorro-Atalaya, O., Arévalo-Tuesta, J., Balarezo-Mares, D., Gonzáles-Pacheco, A., Mendoza-León, O., Quipuscoa-Silvestre, M., Tomás-Quispe, G. And Suarez-Bazalar, R., 2023. K-fold cross-validation through identification of the opinion classification algorithm for the satisfaction of University Students. *International Journal of Online and Biomedical Engineering (iJOE)*, 19(11).
- C., L., Mark A... Latino, Robert J... Latino, Kenneth, 2021. *Root cause analysis: Improving performance for bottom-line results*, fifth edition. S.I.: CRC PRESS.
- Jo, T., 2019. *Text mining concepts, implementation, and Big Data Challenge*. Cham: Springer International Publishing.
- Julianto, Y., Setiabudi, D.H. And Rostianingsih, S., 2022. *Analisis Sentimen Ulasan Restoran Menggunakan Metode Support Vector Machine*.
- Kaushik, K., Bhardwaj, A., Dahiya, S., Maashi, M.S., Al Moteri, M., Aljebreen, M. And Bharany, S., 2022. Multinomial naive bayesian classifier framework for systematic analysis of smart IOT devices. *Sensors*, 22(19), p.7318.
- Maulidah, I., Widodo, J. And Zulianto, M., 2019. Pengaruh Kualitas Produk Dan Kualitas Pelayanan Terhadap Kepuasan Konsumen Di Rumah Makan Ayam Goreng Nelongso Jember. *Jurnal Pendidikan Ekonomi: Jurnal Ilmiah Ilmu Pendidikan, Ilmu Ekonomi dan Ilmu Sosial*, 13(1), p.26. <https://doi.org/10.19184/jpe.v13i1.10416>.
- Simangunsong, J., Zarlis, M. And Tulus, 2019. Analysis of algorithms support vector machine with naive Bayes kernel in Data Classification. *Proceedings of the International Conference on Natural Resources and Technology*.
- Park, S.M. And Kim, Y.G., 2021. Root Cause Analysis Based on Relations Among Sentiment Words. *Cognitive Computation*, 13(4), pp.903–918. <https://doi.org/10.1007/s12559-021-09872-3>.
- Raschka, S., Patterson, J. And Nolet, C., 2020. *Machine learning in Python: Main developments and technology trends in Data Science, Machine Learning, and Artificial Intelligence*. *Information*, 11(4), p.193.
- Rofiqoh, U., Setya Perdana, R. And Fauzi, M.A., 2017. *Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexicon Based Features*.
- Ruhyana, N., 2019. *ANALISIS SENTIMEN TERHADAP PENERAPAN SISTEM PLAT NOMOR GANJIL/GENAP PADA TWITTER DENGAN METODE KLASIFIKASI NAIVE BAYES*.
- Sifa Amalia, B., Umaidah, Y., Mayasari, R., Karawang Jl HSRonggo Waluyo, S., Telukjambe Timur, K. and Karawang, K., 2021. *ANALISIS SENTIMEN REVIEW PELANGGAN RESTORAN MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN K-NEAREST NEIGHBOR*. 19(1), pp.28–34.

*Halaman ini sengaja dikosongkan*