Abimanyu, 2025
https://www.ojs.unr.ac.id/index.php/Equilibrium
E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

MBE: JURNALMANAJEMEN DAN BISNIS EQUILIBRIUM

SENTIMENT ANALYSIS OF CUSTOMER SATISFACTION TOWARDS STEAK HUT MANYAR KERTOARJO RESTAURANT SERVICES USING THE TF-IDF METHOD

Rama Chikal Abimanyu[1],Istian Kriya Almanfaluti[2]
[1,2]Universitas Muhammadiyah Sidoarjo, Jawa Timur , Indonesia
rama.abimanyu@gmail.com[1], istian.alman@umsida.ac.id[2]

Abstract: On online shopping sites or often referred to as marketplaces, there is a column of comments and reviews of transactions that have been made by buyers for products that have been purchased. With this product assessment feature, buyers can consider decisions about the products they will buy. But at this time there is a problem with the review feature because many buyers give negative comments but give a five-star rating. This results in the feature of giving values from consumers being bad. For this reason, a sentiment analysis study was conducted on the review feature at the Steakhut Manyar restaurant using the naive Bayes method and the Tf-Idf algorithm. Based on the review of reviews at the Steakhut restaurant, 1000 review data have been collected which are divided into two, namely 700 training data and 300 test data. After that, the text preprocessing data stage is carried out, where the text preprocessing stage is collecting product and service review data on the web page (Cleaning data), changing uppercase letters to lowercase letters (Casefolding), separating sentences into single sentences (tokenizing), removing conjunctions that are not used for sentiment analysis (stopwords), changing words to basic words (stemming) and continuing to give weight to each word using the Tf-idf algorithm.

Keywords: *TF-Idf, Sentimen Analysis, Naïve Bayes*

INTRODUCTION

In the current digital era, digital platforms are frequently used by the public to find restaurants with the best product and service quality. One such prominent platform is TripAdvisor (Kosasih & Alberto, 2021). For culinary tourism, the quality of service and food significantly impacts customer satisfaction and loyalty (Mastan & Toni, 2020). Tripadvisor serves as a benchmark for restaurant popularity, complemented by consumer reviews. Therefore, customer satisfaction with service and food can be analyzed from comments found on digital platforms like TripAdvisor, often utilizing techniques such as text preprocessing (Azzahra & Wibowo, 2020).

The reviews and ratings provided by TripAdvisor are crucial for customers to evaluate service quality and make informed choices. These reviews generally act as a primary parameter for consumers. Consequently, TripAdvisor offers a review feature that helps restaurants build strong brand awareness and enables consumers to find better dining options (Sari et al., 2019). Nearly 30% of internet users consistently provide online

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

comments or reviews on products or services. These consumer comments often pertain to service quality, product quality, and environmental quality (Neri et al., 2012).

Good service quality, such as being friendly to consumers and providing a comfortable environment that meets customer needs, significantly influences customer satisfaction. Excellent service is as vital as food or product quality, as every customer naturally desires the best possible service(Maulidah et al., 2019). Sentiment analysis itself is a depiction of the polarity of text or words. entiment analysis of restaurant visitors' opinions regarding service and food is crucial for generating recommendations that help restaurant owners deliver product and service quality aligned with visitor preferences(Singgalen, 2023). Search results on TripAdvisor typically display a list of suitable restaurants, often showing images, restaurant names, addresses, overall food and service reviews, and a hyperlink to the restaurant's profile page.

Currently, the Big Data era allows business developers to identify and analyze consumer preferences for restaurant service and food (Fauzi et al., 2021). The algorithm used for sentiment analysis of review data is critical in determining the accuracy, precision, and sensitivity that serve as performance benchmarks for the chosen algorithm(Sari et al., 2019). However, a common challenge arises from the inconsistency between the written reviews and the given ratings for products, making it difficult for buyers to discern product quality. This inconsistency also significantly affects the accuracy of restaurant comments and ratings due to the numerous discrepancies between review content and star ratings.

LITERATURE REVIEW

Tripadvisor

TripAdvisor is a website that offers online reviews detailing travel experiences and visits to restaurants, hotels, and accommodations(Barreda & Bilgihan, 2013). It provides a vast directory of restaurants and hotels that can be recommended for visits, both nationally and internationally.

Digital Marketing

Digital marketing is a strategy used by businesses or SMEs to promote their products using technology like websites, blogs, and social media advertising(Pike & Page, 2014). It helps introduce or market products to consumers with a wider reach, leveraging technological advancements

Consumer

A consumer is an individual or group who engages in a purchase transaction for goods or services from a seller(Hidayani,2008). Consumers are also active buyers on marketplace platforms and are

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

often the primary target for sellers promoting their products. Essentially, a consumer is an individual or group who consumes a seller's products, whether they are edible goods or services.

Customer satisfaction

Customer satisfaction relates to a specific interest and often involves the positive spread of that interest from one individual to another(Kim et al., 2009). Customer satisfaction consistently influences consumer perceptions of a restaurant's service and food quality, frequently leading to recommendations for future visits to others.

Service Quality

Service quality refers to the interaction between an individual as an employee and their proficient and professional service capabilities at the workplace. It reflects the value customers place on a restaurant concerning the quality of its service(Liu & Lee, 2016). Excellent service quality significantly contributes to building positive brand awareness.

Analysis Sentiment

Sentiment analysis is a study of text production related to the broader field of specific data processing activities(Basari et al., 2013). It is commonly used to gauge the accuracy of reviews on marketplace and e-commerce platforms, aiming to understand how precise the platform's ratings are for a particular outlet.

Brand Awareness

Brand awareness is the ability of prospective buyers to recall and recognize a specific brand for a particular product(Apriani et al., 2019). It greatly impacts sales because a brand is the most recognizable name to consumers. Whether a positive or negative incident occurs, brand awareness helps businesses ensure their consumers consistently remember their products.

Metode Naive Bayes

The Naive Bayes calculation is a statistical classification method used to determine the probability of an item belonging to a specific class (Valkanas et al., 2013).The Naive Bayes method is applied after text weighting, typically performed using the TF-IDF algorithm.

Algoritma *Tf-Idf*

TF-IDF is a comparative method used for word weighting. It calculates the frequency of a word's occurrence within a document and the inverse document frequency, which compares the word's prevalence across multiple documents (Andriani & Wibowo, 2021).

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

METHODS

This section will discuss the research stages to be conducted for sentiment analysis of food products found on the TripAdvisor website. Sentiment analysis is a useful method for classifying reviews as negative or positive(Shathik & Prasad, 2020). The method used in this research is Naive Bayes, utilizing the TF-IDF algorithm on product reviews for Steak Hut from the TripAdvisor e-marketplace. The term "Naive" in Naive Bayes indicates that when an element appears, it has no direct connection to other features; thus, each element contributes independently to the classification without depending on other features. The reason for choosing the Naive Bayes method is its higher accuracy, even with a smaller data setting (Apriani & Gustian, 2019).

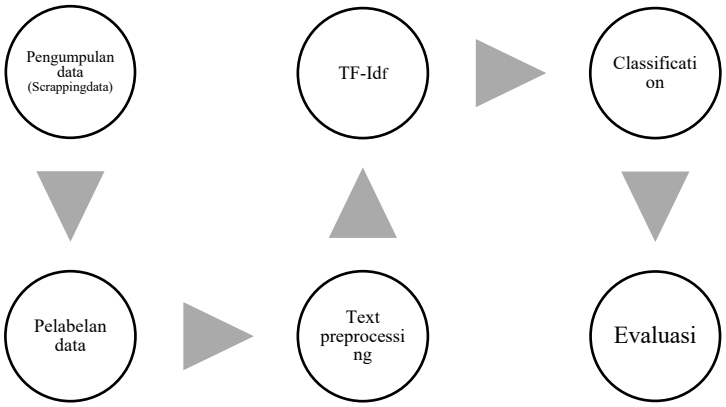
Research Object

The research object for this article is the review data of Steak Hut restaurant obtained from the TripAdvisor website. This study focuses on the sentiment analysis of consumer reviews given to the restaurant on TripAdvisor. The research aims to evaluate customer satisfaction and provide recommendations for improvements to the restaurant. The data analyzed comprises 1,000 review entries, which will be classified using the Naive Bayes method and weighted using the TF-IDF algorithm..

Research Model

A research model is a fundamental concept derived from reality, observation, and literature reviews. The following outlines the steps that will be performed and implemented, as illustrated in Figure 1 (referring to an image that was not provided in the prompt).

Figure 1.
Research Model



Source Research Model (Tazkia & Arkhiansyah, n.d.2024)

DOI: 10.47329/jurnal_mbe.v11i2.1463

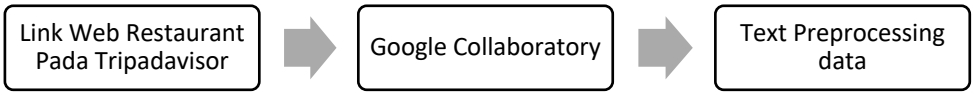
E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

Information:

Data Collecting (Scrapping Data)

Data scraping is the process of collecting or extracting data from a source for research purposes(Valkanas et al., 2013). It also refers to the automated collection of data and information from websites.In this stage, 1,000 data points consisting of product and service reviews from restaurants on the TripAdvisor website will be collected. These 1,000 data points will then be divided into two sets: 700 for training data and 300 for testing data. Data collection will employ data selection, which serves as a more effective and efficient tool for determining the amount of training data .An example of data selection usage is to remove data not required for the classification process After data acquisition,the After the data collection process, we can carry out several stages, we can see these stages in the figure below..

Figure 2.
Scrapping data processing



Source Research Model (Tazkia & Arkhiansyah, n.d.2024)

Review Data Labelling

In this step, the collected data from the website will be dissected and categorized into two labels: positive and negative. Once the review data has been labeled, it will undergo a preprocessing data phase where the data will be processed into a more effective dataset. Data labeling is highly beneficial for the preprocessing data stage.

Data Preprocessing

Common text preprocessing techniques include pre-processing, case folding (cleaning), tokenization, normalization, stopword removal, and stemming(Najjichah et al., 2019). Data preprocessing is the stage where the text reviews are standardized in terms of their form and format, after which they proceed to the next process. This is done to improve data quality, prevent issues during data processing, and make the dataset more effective and efficient. In this stage, consumer reviews or raw data are transformed into a format that can be processed for subsequent steps.

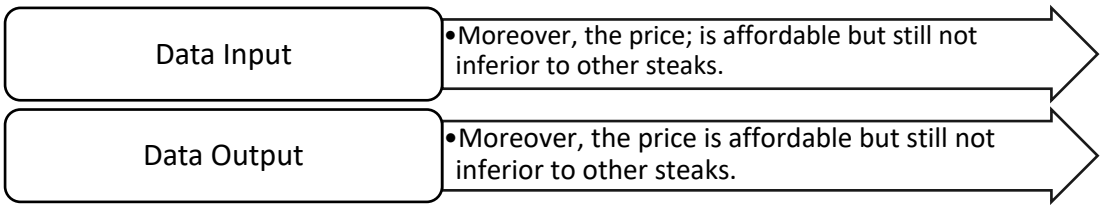
DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

Cleaning

Cleaning involves removing all unnecessary attributes from the text reviews, such as hashtags, symbols, punctuation, numbers, and URLs. This makes the data more efficient and improves its processability. An example of the cleaning process can be seen in figure 3.

Figure 3.
Example data cleaning result

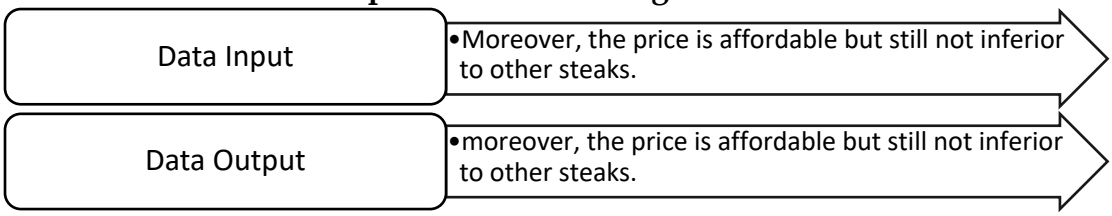


Source Data processing results, 2025

Case Folding

Case folding is the process of converting all uppercase letters in the text reviews to lowercase. The purpose is to standardize the format of the review document. An example of the case folding process can be seen in Figure 4.

Figure 4.
Example data case folding result

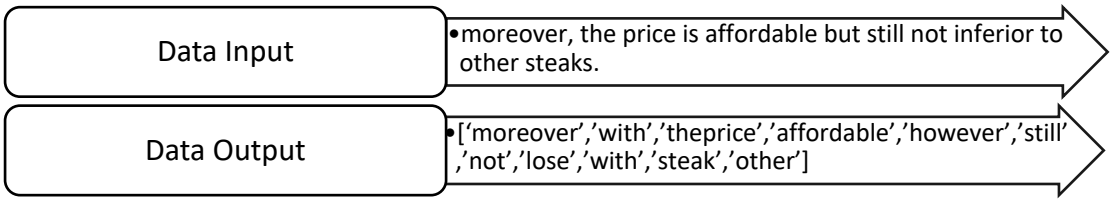


Source Data processing results, 2025

Normalisasi

Normalization is the process of transforming incomprehensible sentences, such as abbreviated sentences, sentences using foreign languages, and non-standard terms, into a standard format. An example of the normalization process can be seen in Figure 5.

Figure 5.
example data normalization result



DOI: 10.47329/jurnal_mbe.v11i2.1463

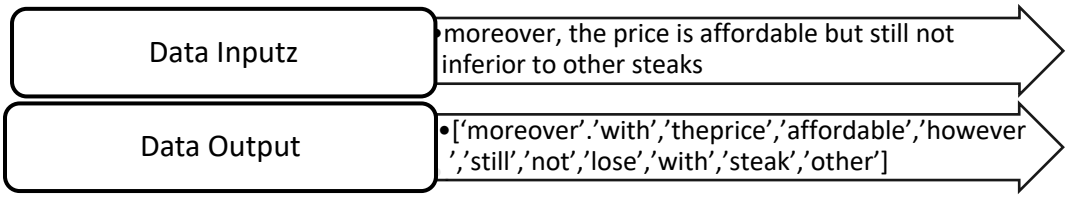
E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

Source Data processing results, 2025

Tokenizing

Tokenizing is the stage of breaking down text into its smallest units or tokens. This facilitates subsequent data processing and analysis steps. An example of the tokenizing process can be seen in Figure 6.

Figure 6.
example data tokenizing result

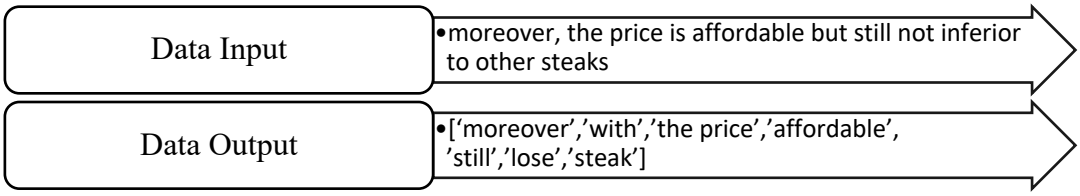


Source Data processing results, 2025

Stopword Removal

Stopword removal is the step of deleting and eliminating words that lack meaning and are useless in the data preprocessing stage. An example of the stopword removal process can be seen in Figure 7.

Figure 7.
example data stopword result

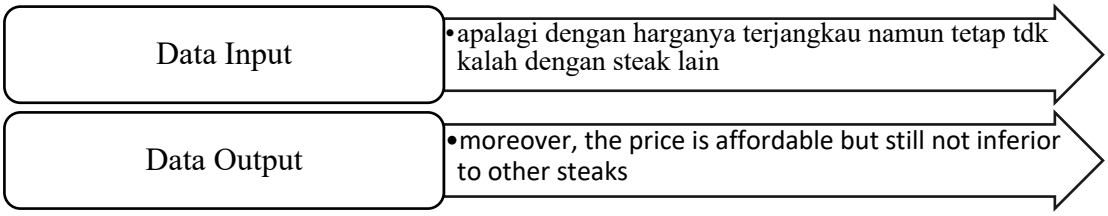


Source Data processing results, 2025

Stemming

Stemming is the process of returning a word to its base form or removing affixes (both prefixes and suffixes) from processed words. An example of the stemming process can be seen in Figure 8.

Figure 8.
example data stemming result



Source Data processing results, 2025

TF-IDF

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

Direct classification of original documents is not possible, hence the need for weighting, which aims to find the weight of words appearing in a collection of documents. The word weighting method is based on a combination. TF-IDF is a comparative method for the word weighting process by calculating the frequency of a word's occurrence within a document and the inverse document frequency that compares it.The TF-IDF algorithm will be used to weight words in the Steak Hut restaurant reviews on the TripAdvisor website to determine the weight of each word in the review text. TF-IDF has a formula for its calculation, as follows:

formula

$$TF - IDF(d,t)=TD(d,t) \times IDF(t) \qquad (1)$$

Information:

Tf : number of times a word appears in a document
IDF: inverse value of documents containing that word
t : word
d : document

Classification

Classification is performed by classifying data using the Naive Bayes algorithm. Naive Bayes calculation is a statistical classification that can be used for the probability of class membership. The Naive Bayes method can also be described as a simple probabilistic classification stage, known for its high processing speed and assumption of independence between features within words. The classification stage requires review data to be used for the Naive Bayes method's classification process. The classification stage requires review data that will be used for the naive Bayes method classification process.

Formula

$$P(A \mid B)= \frac{P(A)\,P(B|A)}{P(B)} \qquad (2)$$

Information:
B=Entered class data
A=Hypothesis data
P(A | B)=Probability of hypothesis A based on condition B.
P(A)=Probability of hypothesis A

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

P(B|A)=Probability of hypothesis B based on condition A
P(B)=Probability of hypothesis B.

Evaluation

The evaluation stage involves calculations including accuracy, precision, recall, and F1-score, performed using a Confusion Matrix on the reviews that have been classified through the Naive Bayes algorithm. The purpose of this analysis is to evaluate the accuracy of the documents after being classified into positive and negative sentiments. In this research, the evaluation process utilizes a confusion matrix. A confusion matrix is a useful tool for assessing the performance of a classification method. The confusion matrix calculation can be seen in Table 1.

Table 1.
Confusion Matrik

Fakta	Prediksi	
	Negative	**Positive**
Negatif	True Negative	False Positive
Positif	False Negative	True Positive

Source Data processing results, 2025

To assess the performance of the Naive Bayes classification, calculations for accuracy, precision, recall, and F1-Score are used, with the following formulas:
Accuracy is the calculation of the number of accurate estimations in a dataset taken from the confusion matrix table. Accuracy can be defined using formula (3):

formula

$$Akurasi=(TN+TP) / (TP + TN + FN + FP) \quad X100\% \qquad (3)$$

Precision is the calculation of the prediction ratio between the number of true positives and false positives. This calculation can be done using formula (4):

formula

$$Presisi= TP/(TP+FP) \ x\ 100\% \qquad (4)$$

Recall is the prediction ratio calculated by dividing the number of true positives by the sum of true positives and false negatives. This calculation can be done using formula (5):

formula

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

$$Recall = TP/(TP+FN) \times 100\% \quad (5)$$

The F1-Score calculation is used to balance precision and recall values. It can be calculated using equation (6):

formula

$$F1\text{-}Score = 2 \times (recall \times presisi) / (recall + presisi) \quad (6)$$

Information:
TP=True Positive
TN=True Negative
FP=False Positive
FN=False Negative

RESULTS AND DISCUSSION

Data Set

After scraping data from the TripAdvisor webpage of @steakhut manyar kertoarjo using a web scraper tool, the collected data consists of customer comments. This data was then compiled into a dataset, totaling 1,000 entries retrieved from the TripAdvisor website. After the data labeling process, all 1,000 data points will be used, as this research focuses solely on positive and negative sentiment comments.

Labelling data

In this research, manual labeling will be performed by a designated labeler. The labeler will assign two categories: a "positive" label for positive sentiments and a "negative" label for negative sentiments. Several comments from the Steak Hut webpage that have already been labeled are shown in Table 2.

Table 2.
Labelling data process

Comment	**lable**
The food in this restaurant is very delicious! I tried their special menu and it was amazing. The service is also fast and friendly.	Positive
My experience at this restaurant was disappointing. The food I ordered arrived late and was not up to expectations. 👎.	Negative
The best place to eat dinner in sby @steakhut. Open until 12 midnight so it's good for hanging out. The price is also affordable.	Positive

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

	Not tasty, because the flour smells like rancid. Spacious place, but slow service. In terms of taste, it's not special	Negative
	I feel the price offered is not commensurate with the quality of the food. I will not return to this restaurant.	Negative

Source Labeling data (Tazkia & Arkhiansyah, n.d.2024)

Preprocessing Data

In the preprocessing stage, text will be classified and cleaned before document analysis is performed. This research utilizes the text preprocessing features available in the NLTK (Natural Language Toolkit) library. The following steps are carried out:

Cleaning Data

During the data cleaning stage, attributes, hashtags, or symbols will be removed from the review data. An example of this process can be seen in Table 3.

Table 3.
Cleaning Data Process

Username	Date	Steakhut Review	Sentiment	Cleaning
Diana	20/7/17	The food in this restaurant is very delicious! I tried their special menu and it was amazing. The service is also fast and friendly.	Positive	The food in this restaurant is very delicious I tried their special menu and it was amazing. The service is also fast and friendly.
Budi	9/1/19	My experience at this restaurant was disappointing. The food I ordered arrived late and was not up to expectations👎.	Negative	My experience at this restaurant was disappointing. The food I ordered arrived late and was not up to expectations.
Lucas	1/15/20	The best place to eat dinner in sby @steakhut. Open until 12 midnight so it's good for hanging out. The price is also affordable	Positive	The best place to eat dinner in sby steakhut. Open until 12 midnight so it's good for hanging out. The price is also affordable

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

Indah	8/2/22	Not tasty, because the flour smells like rancid. Spacious place, but slow service. In terms of taste, it's not special.	Negative	Not tasty because the flour smells like rancid. Spacious place, but slow service. In terms of taste it's not special
Halimah	13/3/22	Saya merasa harga yang ditawarkan tidak sebanding dengan kualitas makanan. Saya tidak akan kembali ke restoran ini.	Negative	I feel the price offered is not commensurate with the quality of the food I will not return to this restaurant

Source Data processing results, (2025)

Case Folding

In the case folding stage, the dataset that has undergone cleaning will be processed to convert all uppercase letters to lowercase (non-capital). You can see an example of the case folding process in Table 4.

Table 4.
Process Case Folding Result

Username	Date	Steakhut Review	Cleaning	Case Folding
Diana	20/7/17	The food in this restaurant is very delicious! I tried their special menu and it was amazing. The service is also fast and friendly.	The food in this restaurant is very delicious I tried their special menu and it was amazing The service is also fast and friendly	the food in this restaurant is very delicious! i tried their special menu and it was amazing the service is also fast and friendly
Budi	9/1/19	My experience at this restaurant was disappointing. The food I ordered arrived late and was not up to expectations👎🏻.	My experience at this restaurant was disappointing. The food I ordered arrived late and was not up to expectations	my experience at this restaurant was disappointing. the food I ordered arrived late and was not up to expectations

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

Lucas	1/15/20	The best place to eat dinner in sby @steakhut. Open until 12 midnight so it's good for hanging out. The price is also affordable	The best place to eat dinner in sby steakhut. Open until 12 midnight so it's good for hanging out. The price is also affordable	the best place to eat dinner in sby steakhut. open until 12 midnight so it's good for hanging out. the price is also affordable
Indah	8/2/22	Not tasty, because the flour smells like rancid. Spacious place, but slow service. In terms of taste, it's not special	Not tasty because the flour smells like rancid. Spacious place but slow service In terms of taste it's not special	not tasty because the flour smells like rancid. spacious place but slow service In terms of taste it's not special
Halimah	13/3/22	I feel the price offered is not commensurate with the quality of the food. I will not return to this restaurant.	I feel the price offered is not commensurate with the quality of the food I will not return to this restaurant	i feel the price offered is not commensurate with the quality of the food i will not return to this restaurant

Source Data processing results, 2025

Normalisasi

In the normalization stage, language that is difficult to understand within the dataset, such as abbreviations, foreign languages, and non-standard language, will be transformed. The results of this normalization process can be seen in Table 5.

Table 5.
Normalisasi Result

Username	Case Folding	Normalisasi
Diana	the food in this restaurant is very delicious! i tried their special menu and it was amazing the service is also fast and friendly	the food in this restaurant is very delicious i tried their special menu and it tastes amazing the service is also fast and friendly
Budi	my experience at this restaurant was disappointing. the food I ordered arrived late and	My experience at this restaurant was disappointing, the food I ordered arrived late and was not as expected

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

	was not up to expectations	
Lucas	the best place to eat dinner in sby steakhut. open until 12 midnight so it's good for hanging out. the price is also affordable	the best place to eat dinner in sby steakhut is open until 12 midnight so it's good for hanging out the prices are also affordable
Indah	not tasty because the flour smells like rancid. spacious place but slow service In terms of taste it's not special	not tasty because the flour smells rancid, spacious place but slow service, taste is not special
Halimah	i feel the price offered is not commensurate with the quality of the food i will not return to this restaurant	I feel the price offered is not commensurate with the quality of the food. I will not return to this restaurant.

Source Data processing results, 2025

Tokenizing

In the tokenizing stage, sentences are broken down into individual words that comprise them. The results of this tokenizing process can be seen in Table 6.

Table 6.
Proses Tokenizing Result

Username	Normalisasi	Tokenisasi
Diana	the food in this restaurant is very delicious i tried their special menu and it tastes amazing the service is also fast and friendly	[food, in, this, restaurant, is, very, delicious, I, tried, their, special, menu, and, the, taste, is, amazing, the, service, is, also, fast, and, friendly]
Budi	My experience at this restaurant was disappointing, the food I ordered arrived late and was not as expected	[my,experience,at,this,restaurant,was,disappointing,the,food,that,I,ordered,came,late,and,did not,match,my,expectations]
Lucas	the best place to eat dinner in sby steakhut is open until 12 midnight so it's good for hanging out the prices are also affordable	[the,best,place,to,eat,dinner,in,sby,steakhut,open,until,12,pm,so,good,for,hangout,price,also,affordable]
Indah	not tasty because the flour smells rancid, spacious place but slow service, taste is not special	[not,nice,because,the,flour,smells,rancid,the,place,is,spacious,but,the,service,is,slow,in,terms,of,taste,not,special]
Halimah	I feel the price offered is	[I feel that the prices offered are not

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

	not commensurate with the quality of the food. I will not return to this restaurant.	commensurate with the quality of the food. I will not return to this restaurant]

Source Data processing results, 2025

Stopword Removal

In the stopword removal stage, words contained in the stopword list will be removed from the dataset. For this purpose, the researcher will use the Indonesian stopword list available in the NLTK (Natural Language Toolkit) library. The stopword removal process can be seen in Table 7.

Table 7.
Stopword Removal Result

Username	Tokenization	Stopword Removal
Diana	[food, in, this, restaurant, is, very, delicious, I, tried, their, special, menu, and, the, taste, is, amazing, the, service, is, also, fast, and, friendly]	[food,restaurant,very, delicious,try,menu,special,taste,excellent,excellent,service,fast,and,friendly]
Budi	[my,experience,at,this,restaurant, was,disappointing,the,food,that,I, ordered,came,late,and,did not,match,my,expectations]	[experience,at,restaurant,disappointing,food,I,ordered,came,late, not,as,expected]
Lucas	[the,best,place,to,eat,dinner,in,sby ,steakhut,open,until,12,pm,so,good,for,hangout,price,also,affordable]	[the,best,place,to,eat, dinner,in,sby,steakhut,open,until,12,night,so,it's,nice,for,hanging, at,affordable,prices]
Indah	[not,nice,because,the,flour,smells, rancid,the,place,is,spacious,but,the,service,is,slow,in,terms,of,taste, not,special]	[delicious,flour,smell, rancid,place,spacious, service,long,about,taste,special]
Halimah	[I feel that the prices offered are not commensurate with the quality of the food. I will not return to this restaurant]	[feel,price,offered,comparable,quality,food,will,not,return,restaurant]

Source Data processing results, 2025

Stemming

In the stemming stage, a word is returned to its root form, or affixes (prefixes or suffixes) are removed from the word. You can see an example of the stemming process in Table 8.

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

Table 8.
Process Stemming Result

Username	Stopword Removal	Stemming
Diana	[food,restaurant,very,delicious,try,menu,special,taste,excellent,excellent,service,fast,and,friendly]	eat at a very delicious restaurant try the special menu the taste is extraordinary the service is fast and friendly
Budi	[experience,at,restaurant,disappointing,food,I,ordered,came,late,not,as,expected]	nature in the restaurant disappointed my food order came late not as expected
Lucas	[the,best,place,to,eat,dinner,in,sby,steakhut,open,until,12,night,so,it's,nice,for,hanging,at,affordable,prices]	the best place to eat dinner in sby steakhut is open until 12 midnight so it's good for hanging out the prices are affordable
Indah	[delicious,flour,smell,rancid,place,spacious,service,long,about,taste,special]	delicious rancid smelling flour, spacious place, long service time, special taste
Halimah	[feel,price,offered,comparable,quality,food,will,not,return,restaurant]	taste price bargain compare quality food not return restaurant

Source Data processing results, 2025

Pembobotan *Tf-idf*

The dataset, having completed the preprocessing stage, will now proceed to be weighted to prepare it for the Naive Bayes method. The results of this TF-IDF weighting calculation can be seen in Table 9.

Table 9.
Pembobotan Tf-Idf result

	Document	Term	TF	DF	IDF	TF-IDF
0	0	eat	1	551	0.620713	0.620713
1	0	reataurant	1	80	2.550421	2.550421
2	0	delicous	1	40	3.243568	3.243568
3	0	try	1	57	2.889397	2.889397
4	0	list	1	147	1.942015	1.942015

Source Data processing results, 2025

DOI: 10.47329/jurnal_mbe.v11i2.1463

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

Evaluation

The model's performance will be evaluated using Multinomial Naive Bayes after the dataset has undergone preprocessing and TF-IDF weighting. The classification using the Naive Bayes method will yield accuracy, precision, recall, and F1-score values, which will be derived from a confusion matrix, as illustrated in Table 10.

Table 10.
Performance Evaluation Results

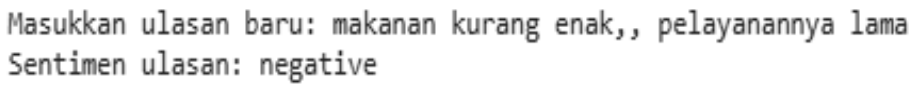
Accuracy : 73.66% Precision : 81.15% Recall : 73.66% F1-Score : 67.37%				
Classification Report :				
	Precision	Recall	F1-Score	Support
Negative	0.72	1.00	0.83	136
Postitive	1.00	0.22	0.36	69
Accuracy			0.74	205
Macro avg	0.86	0.61	0.60	205
Weighted avg	0.81	0.74	0.67	205

Source Data processing results, 2025

Testing

In the testing process, new reviews are randomly entered and then immediately labeled as either negative sentiment or positive sentiment. This can be seen in the image 9 below.

Figure 9.



Results of Testing with New Review Input

Source Data processing results, 2025

CONCLUSION

Based on the results, we can conclude that the Naive Bayes method provides accurate results in determining the accuracy of ratings assigned to consumer reviews for a company, thereby assisting consumers in their evaluations. The Naive Bayes method achieved an accuracy ranging from 73.66% to 81.15%, demonstrating its ability to deliver accurate values and results during both the training and testing phases. However, in the confusion matrix evaluation, the True Negative and False Negative values were suboptimal, failing to achieve perfect scores.

DOI: 10.47329/jurnal_mbe.v11i2.1463

This research focused on sentiment analysis using the Naive Bayes method, a more sophisticated approach for assessing the accuracy of consumer reviews to evaluate a company's performance. It is hoped that future research can achieve even more perfect results in determining the accuracy of sentiment analysis for companies. This, in turn, should further assist consumers in evaluating a company's performance.

.

SUGGESTIONS

First, increasing the size and diversity of the dataset would be beneficial. Gathering a larger volume of reviews from multiple online platforms could help create a more generalized model and reduce bias associated with a single source such as TripAdvisor. This would improve the robustness and applicability of the findings, exploring alternative or hybrid classification methods alongside Naive Bayes could lead to better performance. Methods like support vector machines, ensemble algorithms, or deep learning techniques such as LSTM or transformers could capture more nuanced linguistic patterns and contextual information, thereby increasing accuracy and reliability.

REFERENCES

Andriani, N., & Wibowo, A. (2021). Implementasi Text Mining Klasifikasi Topik Tugas Akhir Mahasiswa Teknik Informatika Menggunakan Pembobotan Tf-Idf Dan Metode Cosine Similarity Berbasis Web. In *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (Senamika) Jakarta-Indonesia*.

Apriani, R., & Gustian, D. (2019). Analisis Sentimen Dengan Naïve Bayes Terhadap Komentar Aplikasi Tokopedia. In *Jurnal Rekayasa Teknologi Nusa Putra* (Vol. 6, Issue 1).

Azzahra, S. A., & Wibowo, A. (2020). *Analisis Sentimen Multi-Aspek Berbasis Konversi Ikon Emosi Dengan Algoritme Naïve Bayes Untuk Ulasan Wisata Kuliner Pada Web Tripadvisor*. *7*(4). Https://Doi.Org/10.25126/Jtiik.202071907

Barreda, A., & Bilgihan, A. (2013). An Analysis Of User-Generated Content For Hotel Experiences. *Journal Of Hospitality And Tourism Technology*, *4*(3), 263–280. Https://Doi.Org/10.1108/Jhtt-01-2013-0001

Basari, A. S. H., Hussin, B., Ananta, I. G. P., & Zeniarja, J. (2013). Opinion Mining Of Movie Review Using Hybrid Method Of Support Vector Machine And Particle Swarm Optimization. *Procedia Engineering*, *53*, 453–462. Https://Doi.Org/10.1016/J.Proeng.2013.02.059

Fauzi, M. R., Pratama, R. A., Laksono, P., & Eosina, P. (2021). Penerapan Big Data Menggunakan Algoritma Multi-Label K-Nearest Neighbor Dalam Analisis Sentimen Konsumen Umkm Sektor Kuliner. *Krea-Tif*, *9*(1), 9. Https://Doi.Org/10.32832/Kreatif.V9i1.3587

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

Hidayani, S. (2008). *Aspek Hukum Perlindungan Konsumen Dalam Pela Yanan Air Bersih Pada Pdam Tirtasari Binjai 0 L E H*.

Kim, W. G., Ng, C. Y. N., & Kim, Y. Soon. (2009). Influence Of Institutional Dineserv On Customer Satisfaction, Return Intention, And Word-Of-Mouth. *International Journal Of Hospitality Management*, *28*(1), 10–17. Https://Doi.Org/10.1016/J.Ijhm.2008.03.005

Kosasih, R., & Alberto, A. (2021). *Analisis Sentimen Produk Permainan Menggunakan Metode Tf-Idf Dan Algoritma K-Nearest Neighbor*. *6*(1). Https://Doi.Org/10.30743/Infotekjar.V6i1.3893

Liu, C. H. S., & Lee, T. (2016). Service Quality And Price Perception Of Service: Influence On Word-Of-Mouth And Revisit Intention. *Journal Of Air Transport Management*, *52*, 42–54. Https://Doi.Org/10.1016/J.Jairtraman.2015.12.007

Mastan, I. A., & Toni, Y. (2020). Analisis Sentimen Terhadap Tempat Kuliner Ayam Gedebuk Dari Komentar Pengunjung Dengan Menggunakan Metode Naïve Bayes Classifier. *Jbase - Journal Of Business And Audit Information Systems*, *3*(1). Https://Doi.Org/10.30813/Jbase.V3i1.2062

Maulidah, I., Widodo, J., & 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), 26. Https://Doi.Org/10.19184/Jpe.V13i1.10416

Najjichah, H., Syukur, A., & Subagyo, H. (2019). Pengaruh Text Preprocessing Dan Kombinasinya Pada Peringkas Dokumen Otomatis Teks Berbahasa Indonesia. In *Jurnal Teknologi Informasi* (Vol. 15, Issue 1). Http://Research.

Neri, F., Aliprandi, C., Capeci, F., Cuadros, M., & By, T. (2012). Sentiment Analysis On Social Media. *Proceedings Of The 2012 Ieee/Acm International Conference On Advances In Social Networks Analysis And Mining, Asonam 2012*, 919–926. Https://Doi.Org/10.1109/Asonam.2012.164

Pike, S., & Page, S. J. (2014). Destination Marketing Organizations And Destination Marketing: Anarrative Analysis Of The Literature. In *Tourism Management* (Vol. 41, Pp. 202–227). Elsevier Ltd. Https://Doi.Org/10.1016/J.Tourman.2013.09.009

Http://Ejournal.Bsi.Ac.Id/Ejurnal/Index.Php/Ji/Article/View/4695

Sari, R. & Nusa, S. (2019). Analisis Sentimen Review Restoran Menggunakan Algoritma Naive Bayes Berbasis Particle Swarm Optimization. *Jurnal Informatika*, *6*(1), 23–28. Http://Ejournal.Bsi.Ac.Id/Ejurnal/Index.Php/Ji/Article/View/4695

Shathik, A., & Prasad, K. (2020). A Literature Review On Application Of Sentiment Analysis Using Machine Learning Techniques. *International Journal Of Applied Engineering And Management Letters (Ijaeml) A Refereed*

E-ISSN: 2723-1704
P-ISSN: 2443-3934
Vol. 11 No. 2 (2025), pp: 135-154

International Journal Of Srinivas University, *4*(2), 2581–7000. Https://Doi.Org/10.5281/Zenodo.3977576

Singgalen, Y. A. (2023). Analisis Sentimen Wisatawan Terhadap Taman Nasional Bunaken Dan Top 10 Hotel Rekomendasi Tripadvisor Menggunakan Algoritma Svm Dan Dt Berbasis Crisp-Dm. *Journal Of Computer System And Informatics (Josyc)*, *4*(2), 367–379. Https://Doi.Org/10.47065/Josyc.V4i2.3092

Tazkia, A., & Arkhiansyah, Y. (N.D.). Implementasi Naive Bayes Classifier Dalam Menganalisis Sentimen Pelanggan Mie Gacoan Pada Instagram. *Ijccs*, *X*, *No.X*, 1–5.

Valkanas, G., Saravanou, A., & Gunopulos, D. (2013). *A Faceted Crawler For The Twitter Service*.

DOI: 10.47329/jurnal_mbe.v11i2.1463