

## Sentiment Analysis of Product Reviews in the Sales System of Putra Electronics Store

Maria Sri Wulandari<sup>1</sup>, Rahayu Noveandini<sup>2</sup>, Yudi Irawan Chandra<sup>3</sup>, Noval Dwi Putra<sup>4</sup>

<sup>1,2,3,4</sup>STMIK Jakarta STI&K, Indonesia

Email: <sup>1</sup>mswuland@yahoo.com, <sup>2</sup>ayu.noveandini@gmail.com,  
<sup>3</sup>yirawanc@gmail.com

### Abstract

The Putra Elektronik Store is a commercial establishment that sells various electronic goods, including mobile phones, desktop computers, and laptops, among other products. The procedure for conducting transactions, which has been ongoing until now, is still done manually. The author suggests developing an online sales platform based on this information. The system design uses UML in addition to the programming languages PHP and Laravel. In addition, a tool for collecting customer reviews has also been developed. The purpose of this feature is to determine whether consumers' feelings regarding the things being offered are favorable or unfavorable by employing the Naive Bayes algorithm and weighting  $tf$  (term frequency) and probability..

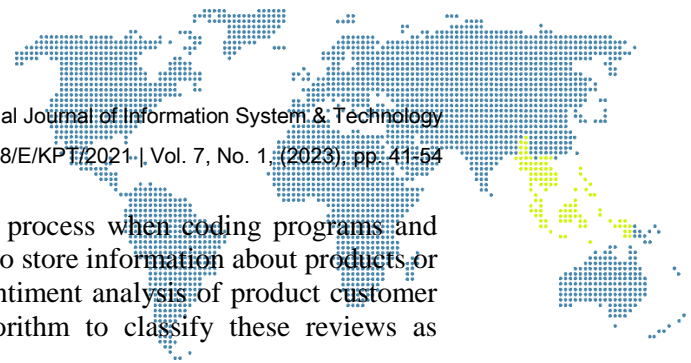
**Keywords:** Online Sales, Sentiment Analysis, Naïve Bayes Algorithm, Term Frequency, Likelihood

### 1. Introduction

Information technology is developing quickly along with the growing human needs, including the Internet. The Internet is a means of information and communication that is fast and accurate. This makes many parties utilize Internet media for various purposes, one of which is for business purposes. Starting from small entrepreneurs to large companies competing to take advantage of advances in Internet technology as a medium that serves to promote products or advertisements via the Internet. Toko Putra Elektronik sells electronics such as cell phones, computers, laptops, etc. Toko Putra Elektronik has several obstacles, namely product marketing activities and the transaction process still uses the old method, which requires customers to be at the store location so that customers who are far from the store location will not be able to see or find out information about the product the customer wants to get to the transaction process.

The obstacles faced by Toko Putra Elektronik must be addressed immediately because these obstacles have a huge impact on store income. The shop owner wants a system that runs online so that customers can get information on a product and carry out the transaction without having to come to the store location because it can be accessed via a cellphone, computer, or laptop device connected to the internet network. In addition to providing information and online transactions, the owner wants a sentiment analysis feature processed through a review from a customer after purchasing a product to find out how the customer responds to the product purchased.

Based on the background above, the author wants to build an online sales system so that customers can see information from the desired product to the stage of providing reviews about the product purchased. The application includes an estimate of product shipping costs from the store location to the customer's address using a third party. In addition, when customers make the payment process, it will be handled directly using a third party so that customers can make the payment process from any bank. To design the system flow process using UML (Unified et al.) diagram, a visual language for modeling and communicating about a system containing diagrams and supporting text. The programming language used to build the system is PHP language using the Laravel



version 8 framework to speed up the development process when coding programs and easy to use, while the DBMS (Database et al.) used to store information about products or transactions is MySQL [6]. For data processing, sentiment analysis of product customer reviews using the naïve Bayes classification algorithm to classify these reviews as positive or negative categories.

## 2. Research Methodology

### 2.1. Sentiment Analysis

One application of machine learning is used to perform sentiment analysis. Machine learning is a technology that allows a machine to learn on its own without human involvement. Machine learning can be used to analyze and collect data to perform certain tasks [9].

*Sentiment* can be defined as an opinion or view that is based on excessive feelings towards something. The sentiment is usually found in statements and sentences that have opinions. The sentiment is also useful for knowing a person's feelings about a particular topic or object [5]. The task performed by sentiment analysis is to classify the polarity in a text, whether contained in documents, sentences, or aspect-level features, and whether the opinions expressed are positive, negative, or neutral [5]. Opinion mining or also known as sentiment analysis is a process that aims to determine whether the polarity of a collection of written text (documents, sentences, paragraphs, etc.) tends to be positive, negative, or neutral (Kontopoulos et al., 2013) [7].

Opinion Mining is a field of computer science that discusses natural language processing, linguistic computing, and text mining, which aims to analyze a person's judgment, sentiment, emotions, attitudes, opinions, assessments of reviews of certain products, organizations, public figures, services, individuals, events, topics, or activities [8]. Sentiment analysis is defined as a data mining process to review reviews or sentences to predict the emotions of sentences through natural language processing (NLP). Sentiment analysis will result in the classification of text into three phases, namely "Positive," "Negative," or "Neutral" [10].

### 2.2. Text Mining

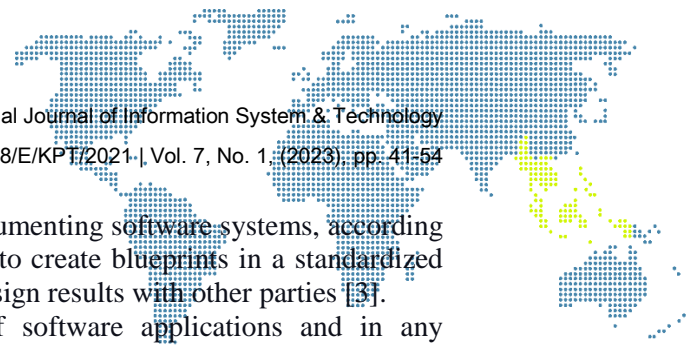
Text mining aims to generate information from a set of documents. Text mining is a solution for retrieving information through processing, grouping, and analyzing large amounts of unstructured data. The process of retrieving information in text mining can produce sentiment analysis that identifies a statement's emotionality, whether positive or negative [4]. Text mining is widely used in various research such as software development, online media, marketing, academics, and politics. Like data mining, text mining will process unstructured data into structured data that is processed using certain algorithms. The process requires an initial stage of text preprocessing for analysis [4].

### 2.3. Preprocessing

Documents generally have an arbitrary or unstructured structure. Therefore, a process is needed to change the previously unstructured form of data into a structured form of data. This conversion process is known as text preprocessing [1]. The following are the stages carried out when performing text preprocessing; cleansing is a process to remove symbols contained in the data; case folding is a process carried out to uniform the characters in the data; filtering is a stage of removing very common words in the data, stemming is a stage that aims to change the term to its root word form, and tokenizing is a stage of cutting strings based on the words that compose them [1].

### 2.4. UML (Unified Modeling Language)

UML (Unified Modeling Language) is a tool that has become a standard in object-oriented software system development. UML is a visual language for completing,



providing specifications, designing models, and documenting software systems, according to Fenny & Esthi. UML allows system developers to create blueprints in a standardized and easy-to-understand form to communicate the design results with other parties [3]. UML can be used as a model for all types of software applications and in any programming language. However, because UML uses classes and operations in its basic concept, UML is more suitable for object-oriented languages such as PHP, C++, Java, C#, or VB.NET [3].

The Waterfall model will be utilized as the development approach for the system, and the following stages will make up the Waterfall model's progression:

1. Analysis: The first step in the design process is to collect data using a direct survey to identify a problem. Next, the system needs that will be prioritized in making the system are analyzed.
2. Design: The system's design includes creating a navigation structure. The system's design using the Unified Modelling Language (UML) includes a class diagram for the database, use case diagrams for process design, activity diagrams, and input-output design to show how the system should look.
3. Implementation: putting the system through its paces by testing utilizing programming languages such as PHP and MYSQL.
4. System Testing: Conducting program experiments using laptops and cellphones with Google Chrome, Internet Explorer, and Mozilla Firefox connected to the internet in a stable connection to predicting potential difficulties or roadblocks to mistakes in the system. The method of testing that can be applied to test Black Box Testing is a type of software testing that examines software without considering the code's or program's internal structure. This type of testing is used to test software.
5. Maintenance: It is necessary to continuously use software and monitor and assess the system's operation continually.

### 2.5. Research stages

Several stages need to be done to facilitate the research and collection of data needed to design and create an online sales system application. These stages include the system design stage and the data collection stage. The system design aims to describe and show a picture of the system to be built. The system design will be implemented into the research using a tool called UML (Unified et al.) which has become a standard for visualizing, designing, and documenting software systems. The following activity diagrams will be applied to the Putra Electronics Store sales system. The activity diagram of the customer registration is in Figure 1.

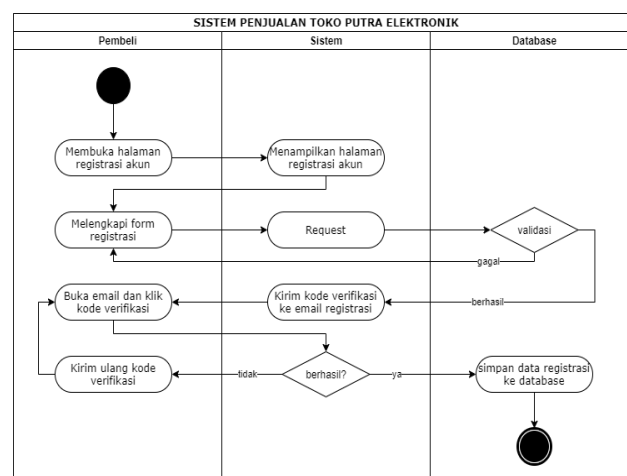
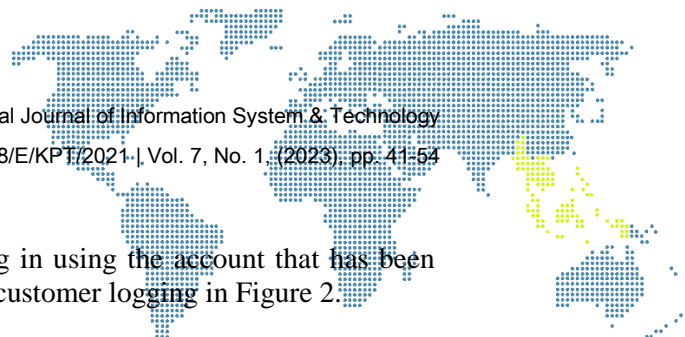
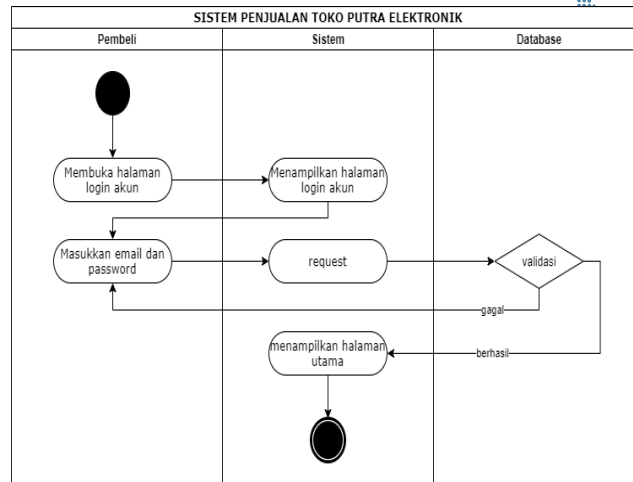


Figure 1. Activity Diagram of Customer Registering

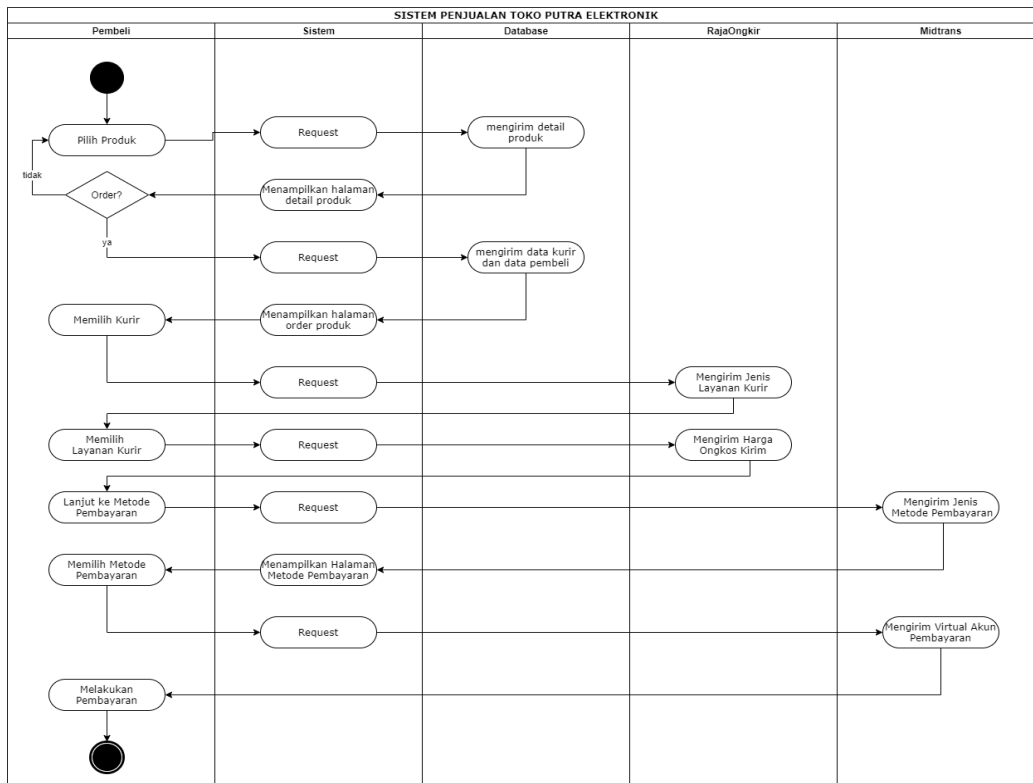


In addition to registering, customers can also log in using the account that has been registered, as depicted in the activity diagram of the customer logging in Figure 2.



**Figure 2.** Activity Diagram of Customer Login

To make an online product order, including the selection of the expedition service you want to use and the desired payment method can be seen in Figure 3.



**Figure 3.** Activity Diagram of Customer Ordering Products

Customers can review products after the product they bought has arrived; customer reviews will be processed using sentiment analysis to determine whether the product has positive or negative reviews. The activity diagram of customers reviewing products is in Figure 4.

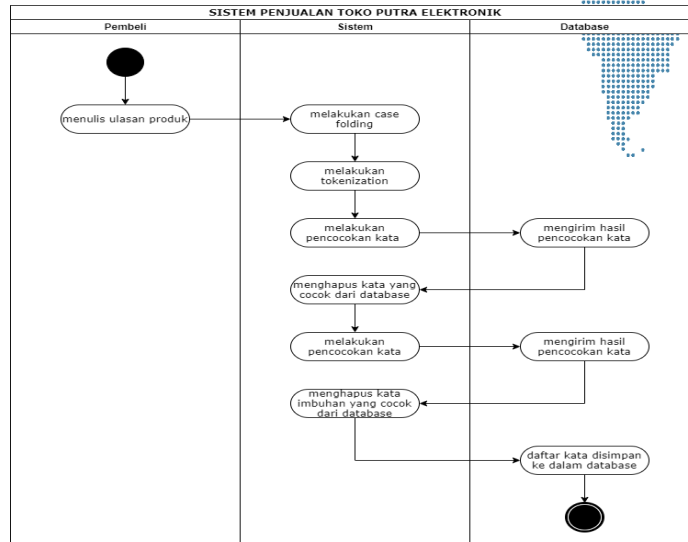
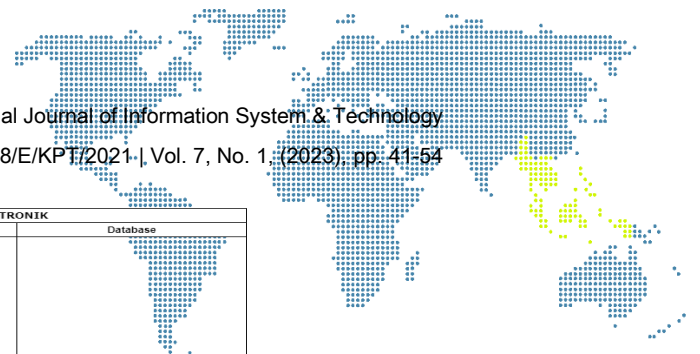


Figure 4. Activity Diagram of Customer Reviewing Products

### 3. Results and Discussions

#### 3.1. Data Collection

The amount of data obtained in the data collection process is 720, consisting of 570 data used as training data which is manually classified into positive and negative according to the nature of the opinion of the data obtained. The following are some examples of training data that have been classified, which can be seen in Table 1.

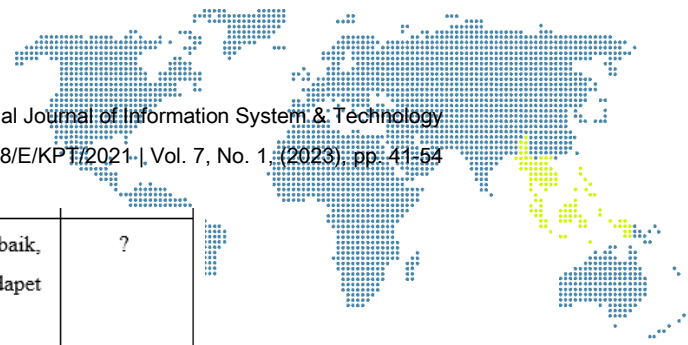
Table 1. Sample Training Data

No.	Kalimat	Kelas
1.	Produknya BAGUSSS BANGET dan berfungsi dengan sangat baik, terima kasih !!	Positif
2.	HALAHHH BARANG MACAM APAAN NIH?? Baru berapa hari pake langsung rusak !	Negatif
3.	Laptopnya keren, modelnya bagus, puas deh sama barangnya!!	Positif
4.	PESAN KAMERA yang datang malah BARANG LAIN, KECEWA PARAH !!!!	Negatif
5.	SSDnya berfungsi dengan baik, kondisinya masih bagus dan mulus	Positif

Meanwhile, test data is obtained from customer reviews when they have purchased a product from the system. The following is an example of test data that will be used for processing on test data, which can be seen in Table 2.

Table 2. Sample Test Data

No.	Kalimat	Kelas
1.	Pas barangnya sampe langsung dipake, alhasil puas banget dah gitu barangnya meskipun murah tapi bagus hasilnya	?



2.	Barangnya berfungsi dengan baik, tapi agak kecewa karena gak dapet kabel casannya, overallpuas sih	?
----	--	---

In addition to the training and test data examples, we need positive and negative root word data to hold the weighting values of the sentences generated from preprocessing. Examples of positive root words are in Table 3.

**Table 3.** Examples of positive root words

No.	Kata Dasar
1.	Bagus
2.	Fungsi
3.	Keren
4.	Puas
5.	Mulus

As for the examples of negative root words used, they are listed in Table 4..

**Table 4.** Examples of Negative Root Words

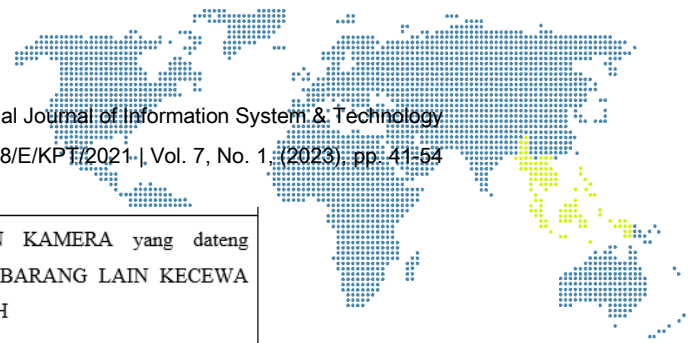
No.	Kata Dasar
1.	Rusak
2.	Kecewa

### 3.2. Preprocessing

Before performing the data classification process using the naïve Bayes algorithm, all training and test data must first pass the preprocessing stage. The preprocessing stage is used to overcome unstructured text documents into structured text documents for further data processing. The first stage is cleansing, which removes all symbols in the data, such as ~!@#\$\$%^&\*()-=\/,. An example of the cleansing process using the training data is in Table 5.

**Table 5.** Cleansing Process Example

No.	Sebelum Cleansing	Sesudah Cleansing
1.	Produknya BAGUSSS BANGET dan berfungsi dengan sangat baik, terima kasih !!	Produknya BAGUSSS BANGET dan berfungsi dengan sangat baik terima kasih
2.	HALAHHH BARANG MACAM APAAN NIH?? Baru berapa hari pake langsung rusak !	HALAHHH BARANG MACAM APAAN NIH baru berapa hari pake langsung rusak
3.	Laptopnya keren, modelnya bagus, puas deh sama barangnya!!	Laptopnya keren modelnya bagus puas deh sama barangnya!!



4.	PESAN KAMERA yang datang malah BARANG LAIN, KECEWA PARAH !!!!	PESAN KAMERA yang datang malah BARANG LAIN KECEWA PARAH
5.	SSDnya berfungsi dengan baik, kondisinya masih bagus dan mulus	SSDnya berfungsi dengan baik kondisinya masih bagus dan mulus

The second stage is case folding, where all letters in a sentence will be converted into lowercase letters, and none of them are capitalized. An example of the case folding process using the training data is in Table 6.

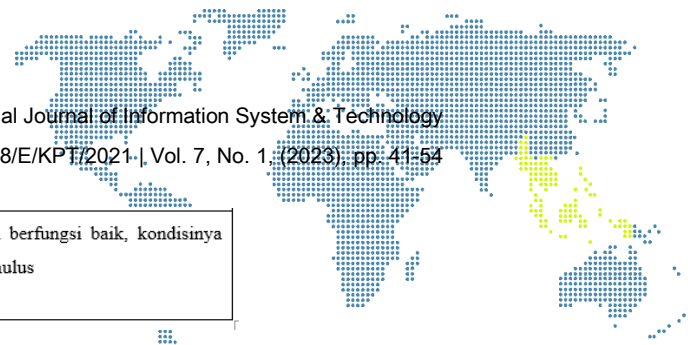
**Table 6.** Example of Case Folding Process

NO	Sebelum Case Folding	Sesudah Case Folding
1.	Produknya BAGUSSS BANGET dan berfungsi dengan sangat baik, terima kasih !!	produknya bagus banget dan berfungsi dengan sangat baik, terima kasih
2.	HALAHHH BARANG MACAM APAAN NIH?? Baru berapa hari pake langsung rusak !	halah barang macam apaan nih?? Baru berapa hari pake langsung rusak !
3.	Laptopnya keren, modelnya bagus, puas deh sama barangnya!!	laptopnya keren, modelnya bagus, puas deh sama barangnya!!
4.	PESAN KAMERA yang datang malah BARANG LAIN, KECEWA PARAH !!!!	pesan kamera yang datang malah barang lain, kecewa parah !!!!
5.	SSDnya berfungsi dengan baik, kondisinya masih bagus dan mulus	ssdnya berfungsi dengan baik, kondisinya masih bagus dan mulus

The third stage is filtering/stopword removal, which is a stage to eliminate words that have no meaning that are feared will interfere with the data classification process. Words that have no meaning, for example, such as "ada", " adalah", "dengan", "dan", "sangat", and so on. An example of the filtering process using training data is in Table 7.

**Table 7.** Example of filtering process

No.	Sebelum Filtering	Sesudah Filtering
1.	Produknya BAGUSSS BANGET dan berfungsi dengan sangat baik, terima kasih !!	Produknya BAGUSS berfungsi
2.	HALAHHH BARANG MACAM APAAN NIH?? Baru berapa hari pake langsung rusak !	HALAHH BARANG pake langsung rusak !
3.	Laptopnya keren, modelnya bagus, puas deh sama barangnya!!	Laptopnya keren, modelnya bagus, puas deh barangnya!!
4.	PESAN KAMERA yang datang malah BARANG LAIN, KECEWA PARAH !!!!	PESAN KAMERA BARANG LAIN, KECEWA PARAH !!!!



5.	SSDnya berfungsi dengan baik,kondisinya masih bagus dan mulus	SSDnya berfungsi baik, kondisinya bagus mulus
----	---	---

The fourth stage is stemming, which is returning every word in a sentence to the basic word by eliminating sentences that have affixes such as "ke-", "-kan", "di-", and so on. An example of the stemming process using training data is in Table 8.

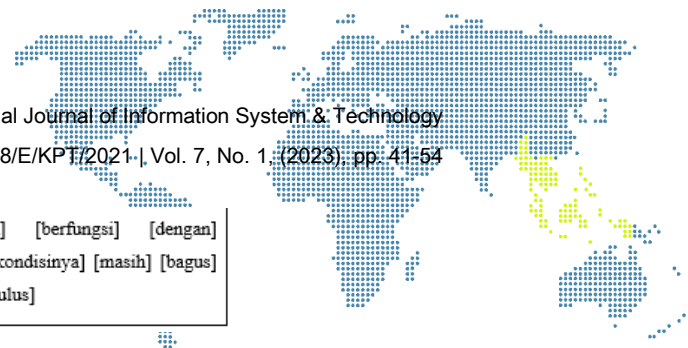
**Table 8.** Example of Stemming Process

No.	Sebelum Stemming	Sesudah Stemming
1.	Produknya BAGUSS BANGET dan berfungsi dengan sangat baik, terima kasih !!	Produk bagus banget dan fungsi dengan sangat baik, terima kasih !!
2.	HALAHH BARANG MACAM APAAN NIH?? Baru berapa hari pake langsung rusak !!	HALAHH BARANG MACAM APA NIH?? Baru berapa hari pake langsung rusak !!
3.	Laptopnya keren, modelnya bagus, puas deh sama barangnya!!	Laptop keren model bagus puas deh sama barang!!
4.	PESAN KAMERA yang datang malah BARANG LAIN, KECEWA PARAH !!!!	PESAN KAMERA yang datang malah BARANG LAIN, KECEWA PARAH !!!!
5.	SSDnya berfungsi dengan baik, kondisinya masih bagus dan mulus	SSD fungsi dengan baik, kondisi masih bagus dan mulus

The fifth stage is tokenizing, which is the stage that breaks the sentence into words that each stand alone so that it is easy to weigh the values. An example of the tokenizing process using training data is in Table 9.

**Table 9.** Example of Tokenizing Process

No.	Sebelum Tokenizing	Sesudah Tokenizing
1.	Produknya BAGUSS BANGET dan berfungsi dengan sangat baik, terima kasih !!	[produknya] [BAGUSS] [BANGET] [dan] [berfungsi] [dengan] [sangat] [baik,] [terima] [kasih] [!!]
2.	HALAHH BARANG MACAM APAAN NIH?? Baru berapa hari pake langsung rusak !!	[HALAHH] [BARANG] [MACAM] [APAAAN] [NIH??] [Baru] [berapa] [hari] [pake] [langsung] [rusak] [!!]
3.	Laptopnya keren, modelnya bagus, puas deh sama barangnya!!	[Laptopnya] [keren,] [modelnya] [bagus,] [puas] [deh] [sama] [barangnya] [!!]
4.	PESAN KAMERA yang datang malah BARANG LAIN, KECEWA PARAH !!!!	[PESAN] [KAMERA] [yang] [datang] [malah] [BARANG] [LAIN,] [KECEWA] [PARAH] [!!!!]



5.	SSDnya berfungsi dengan baik, kondisinya masih bagus dan mulus	[SSDnya] [berfungsi] [dengan] [baik,] [kondisinya] [masih] [bagus] [dan] [mulus]
----	--	--

The following is an example of the training data fully preprocessed from the first to the last stage in Table 10.

**Table 10.** Example of Preprocessing Process

No	Sebelum Preprocessing	Sesudah Preprocessing
1.	Produknya BAGUS BANGET dan berfungsi dengan sangat baik, terima kasih !!	[produknya] [bagus] [banget] [fungsi] [terima] [kasih]
2.	HALAHH BARANG MACAM APAAN NIH?? Baru berapa hari pake langsung rusak !!	[halahh] [barang] [langsung] [rusak]
3.	Laptopnya keren, modelnya bagus, puas deh sama barangnya!!	[laptopnya] [keren] [model] [bagus] [puas] [barang]
4.	PESAN KAMERA yang datang malah BARANG LAIN, KECEWA PARAH !!!!	[pesan] [kamera] [barang] [kecewa] [parah]
5.	SSDnya berfungsi dengan baik, kondisinya masih bagus dan mulus	[ssd] [fungsi] [kondisi] [bagus] [mulus]

After preprocessing, the data will be weighted using tf (term-frequency) to calculate the number of occurrences of negative and positive root words grouped by class from the training data. An example of base word weighting is in Table 11.

**Table 11.** Base Word Weighting

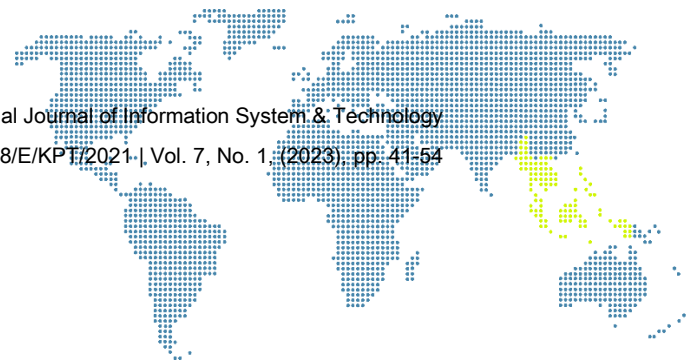
No.	Kata Dasar	Tf (Positif)	Tf (Negatif)
1.	Bagus	3	0
2.	Fungsi	2	0
3.	Rusak	0	1
4.	Keren	1	0
5.	Puas	1	0

### 3.3. Naïve Bayes Classification

After performing the preprocessing stage, which aims to tidy up unstructured sentences into structured sentences, the next step is to perform the calculation process using the naïve Bayes classification algorithm.

#### 1. Naïve Bayes Classification Testing Using Training Data

- a. Calculating the prior probability of each class, there are two classes, namely positive and negative.



$$P(\text{Positif}) = \frac{f_x(\text{positif})}{|C|} = \frac{1}{2} = 0,5$$

$$P(\text{Negatif}) = \frac{f_x(\text{negatif})}{|C|} = \frac{1}{2} = 0,5$$

b. Calculating the probability likelihood of each term in the document. The total number of positive base words and negative base words in Table 11 is 8, where 7 are total positive terms, and 1 is total negative terms. The number of terms depends on the results of sentence preprocessing. To calculate the likelihood probability based on the words in Table 11 using the following formula:

$$P(w|\text{positif/negatif}) = \frac{(nk(\text{positif/negatif}) + 1)}{(n.\text{positif/negatif}) + |T|}$$

Where:

1.  $P(w|\text{positif/negatif})$  is the tested base word with positive or negative class..
2.  $(nk(\text{positif/negatif}))$  is the term value of the base word with a positive or negative class.
3.  $(n.\text{positif/negatif})$  is the total term value of positive and negative classes, where the total positive class term value is 7, and the total negative class term value is 1.
4.  $|T|$  is the total number of positive and negative base words, which is 8.

The following is the calculation of each word in Table 11 using the probability likelihood formula:

1. Likelihood probability of the word "Bagus"

$$P(\text{bagus}|\text{positif}) = \frac{nk(\text{positif}) + 1}{n.\text{positif} + |T|} = \frac{3 + 1}{8 + 7}$$

$$= 0.2666666666666667$$

$$P(\text{bagus}|\text{negatif}) = \frac{nk(\text{negatif}) + 1}{n.\text{negatif} + |T|} = \frac{0 + 1}{2 + 7}$$

$$= 0.1111111111111111$$

2. Likelihood probability of the word "Fungsi"

$$P(\text{fungsi}|\text{positif}) = \frac{nk(\text{positif}) + 1}{n.\text{positif} + |T|} = \frac{2 + 1}{8 + 7}$$

$$= 0.2$$

$$P(\text{fungsi}|\text{negatif}) = \frac{nk(\text{negatif}) + 1}{n.\text{negatif} + |T|} = \frac{0 + 1}{2 + 7}$$

$$= 0.1111111111111111$$

3. Probability likelihood of the word "Rusak"

$$P(\text{rusak}|\text{positif}) = \frac{nk(\text{positif}) + 1}{n.\text{positif} + |T|} = \frac{0 + 1}{8 + 7}$$

$$= 0.0666666666666667$$

$$P(\text{rusak}|\text{negatif}) = \frac{nk(\text{negatif}) + 1}{n.\text{negatif} + |T|} = \frac{1 + 1}{2 + 7}$$

$$= 0.2222222222222222$$

4. Probability likelihood of the word "Keren"

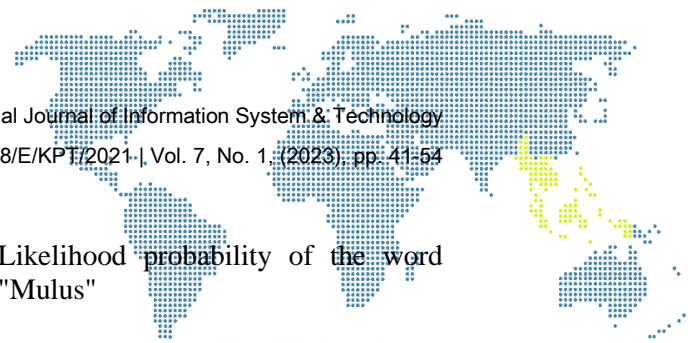
$$P(\text{keren}|\text{positif}) = \frac{nk(\text{positif}) + 1}{n.\text{positif} + |T|} = \frac{1 + 1}{8 + 7}$$

$$= 0.1333333333333333$$

$$P(\text{keren}|\text{negatif}) = \frac{nk(\text{negatif}) + 1}{n.\text{negatif} + |T|} = \frac{1 + 1}{2 + 7}$$

$$= 0.1111111111111111$$

5. Probability likelihood of the word "Puas"



$$P(\text{puas}|\text{positif}) = \frac{nk(\text{positif}) + 1}{n.\text{positif} + |T|} = \frac{1 + 1}{8 + 7}$$

$$= 0.1333333333333333$$

$$P(\text{puas}|\text{negatif}) = \frac{nk(\text{negatif}) + 1}{n.\text{negatif} + |T|} = \frac{0 + 1}{2 + 7}$$

$$= 0.1111111111111111$$

6. Likelihood probability of the word "Kecewa"

$$P(\text{kecewa}|\text{positif}) = \frac{nk(\text{positif}) + 1}{n.\text{positif} + |T|} = \frac{0 + 1}{8 + 7}$$

$$= 0.0666666666666667$$

$$P(\text{kecewa}|\text{negatif}) = \frac{nk(\text{negatif}) + 1}{n.\text{negatif} + |T|} = \frac{1 + 1}{2 + 7}$$

$$= 0.2222222222222222$$

7. Likelihood probability of the word "Mulus"

$$P(\text{mulus}|\text{positif}) = \frac{nk(\text{positif}) + 1}{n.\text{positif} + |T|} = \frac{1 + 1}{8 + 7}$$

$$= 0.1333333333333333$$

$$P(\text{mulus}|\text{negatif}) = \frac{nk(\text{negatif}) + 1}{n.\text{negatif} + |T|} = \frac{0 + 1}{2 + 7}$$

$$= 0.1111111111111111$$

3.4. Naïve Bayes Classification Testing Using Test Data

After knowing the results of preprocessing the test data, perform the test data classification process by multiplying all the opportunity values. The highest result of the multiplication of the opportunity value will be the new class of the data.

Table 12. Test Data Preprocessing Example

No.	Before Preprocessing	After Preprocessing
1.	Pas barangnya sampe langsung dipake, alhasil puas banget dah gitu barangnya meskipun murah tapi bagus hasilnya	[puas] [bagus]
2.	Barangnya berfungsi dengan baik, tapi agak kecewa karena gak dapet kabel casannya, overall puas sih	[fungsi] [kecewa] [puas]

1. In table 12 test data that has been preprocessed, there are the words "puas", and "bagus".

$$P(\text{uji}|\text{positif})$$

$$= P(\text{positif}) \times P(\text{puas}|\text{positif}) \times P(\text{bagus}|\text{positif})$$

$$= 0.5 \times 0.1333333333333333 \times 0.2666666666666667$$

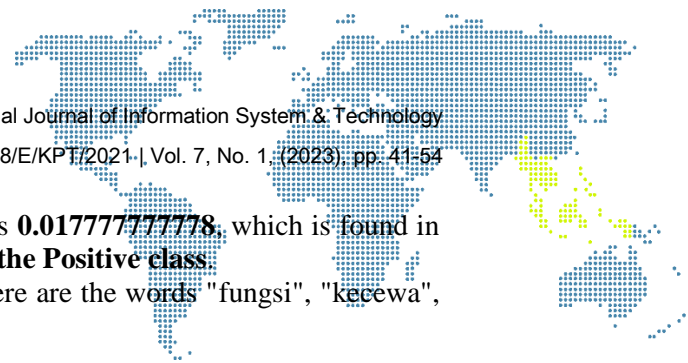
$$= 0.0177777777777778$$

$$P(\text{uji}|\text{negatif})$$

$$= P(\text{negatif}) \times P(\text{puas}|\text{negatif}) \times P(\text{bagus}|\text{negatif})$$

$$= 0.5 \times 0.1111111111111111 \times 0.1111111111111111$$

$$= 0.0061728395061728$$



**Conclusion:** The highest probability value is **0.017777777778**, which is found in  $P(\text{test}/\text{positive})$ , so the data is classified into **the Positive class**.

- In the preprocessed test data of table 12, there are the words "fungsi", "kecewa", and "puas".

$$P(\text{uji}/\text{positif})$$

$$= P(\text{positif}) \times P(\text{fungsi}/\text{positif}) \times P(\text{kecewa}/\text{positif}) \times P(\text{puas}/\text{positif})$$

$$= 0.5 \times 0.2 \times 0.0666666666666667 \times 0.1333333333333333$$

$$= 0.00888888888888889$$

$$P(\text{uji}/\text{negatif})$$

$$= P(\text{negatif}) \times P(\text{fungsi}/\text{negatif}) \times P(\text{kecewa}/\text{negatif}) \times P(\text{puas}/\text{negatif})$$

$$= 0.5 \times 0.1111111111111111 \times 0.2222222222222222 \times 0.1111111111111111$$

$$= 0.0013717421124829$$

**Conclusion:** The highest probability value is **0.0013717421124829**, found in  $P(\text{test}/\text{negative})$ , so the data is classified into **the Negative class**.

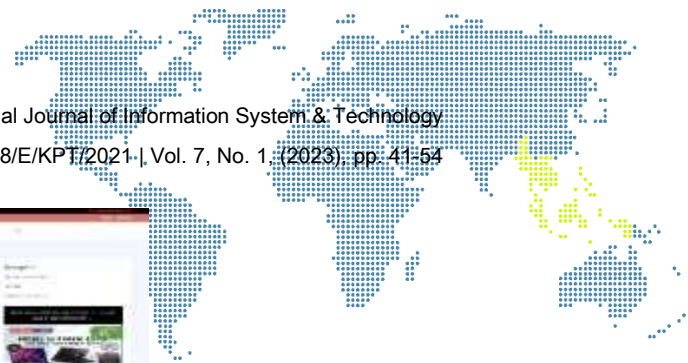
After successfully carrying out the system design stage and the data collection stage, this section will explain how the implementation results after coding the program based on the system design to achieve the desired results in the form of an online sales system so that customers do not have to come to the location to see or find out information about the desired product or make product subscription transactions because this can be done online via cellphone, laptop, or computer as long as it is connected to the internet network.

Here are some views of the sales system that has been successfully created. For the initial page, when first accessed, several displays of the latest products and featured products will appear; the display can be seen in Figure 5.



**Figure 5.** Initial Display of the Sales System

Meanwhile, the page to view complete information about the product starting from the product name, brand, rating, description, specifications, and reviews about the product, can be seen in Figure 6.



**Figure 6.** Product Detail Display

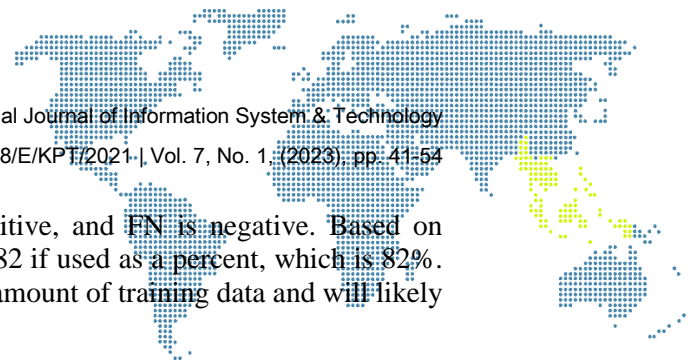
After the customer purchases a product and the customer has received the product, the customer can provide a review of the product that has been purchased. Reviews submitted by customers will be processed using sentiment analysis to classify whether the review is positive or negative. Before getting the classification results, the test data will first carry out the preprocessing stage so that the data becomes clean or does not contain noise to facilitate assessment. The preprocessing stage has several processes, including; cleansing, case folding, filtering/stopword removal, stemming, and tokenizing. After the clean test data from the symbols after the cleansing process, the next step is case folding. Test data that has been case-folded no longer have capital letters because they have been changed to lowercase. The next process is filtering/stopword removal on the test data that has been case folded. Some words that have no meaning, such as "dan," "yang," and "dengan," will be deleted after the filtering process is carried out. The next process is stemming the filtered test data. Data containing affixes such as "ke-," "-nya," and "di-" will be deleted and returned to the basic word form. The next process is tokenizing the stemming test data. When the data has been cleaned through the preprocessing stage, the tokenized data will be processed by sentiment analysis using the Naïve Bayes algorithm in the PHP programming language. The accuracy of the overall classification results is in Table 13.

**Table 13.** Test Data Accuracy Results

Keterangan	True Positif	True Negatif
Prediksi Positif	68	15
Prediksi Negatif	12	55

Table 13 shows that the classification results using the naïve Bayes algorithm get 68 positive data according to predictions, 12 other positive data are classified as negative, and 55 negative data according to predictions, but 15 other data are classified as positive. To calculate the accuracy value, you can use formula 3.

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP + TN}{TP + TN + FP + FN} \\
 &= \frac{68 + 55}{68 + 55 + 12 + 15} \\
 &= \frac{123}{150} = 0,82
 \end{aligned}$$



TP is positive, TN is negative, FP is false positive, and FN is negative. Based on manual calculations using formula 3, the result is 0.82 if used as a percent, which is 82%. The level of accuracy will always increase with the amount of training data and will likely reach 90%.

#### 4. Conclusion

Based on the results of the research and discussion, it can be concluded that customers who want to find information about products or want to buy products from Putra Elektronik Stores can be done online via mobile phones, computers, or laptops that are connected to the internet network so they do not need to take the time to cover the distance to the shop location which results in draining stamina and time for other activities. The sentiment analysis method is also successfully applied when a customer writes a review about a product that has been purchased with an accuracy of 82% so that the store will know what the customer thinks about the product they are selling, whether it produces positive or negative reviews.

#### References

- [1] Jumeilah, F. S., "Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian", *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 1(1), 19-25, 2017.
- [2] Rahmawati, F. H., & Adityarini, E., "Sistem Informasi Persediaan Barang pada CV. Anak Teladan", *Jurnal Sistem Informasi*, 10(1), 1-7, 2017.
- [3] Anardani, S., "Perancangan Sistem Berorientasi Objek Dengan Pemodelan Uml (Unified Modeling Language) Tools", 2019.
- [4] Mesran, O. K. S., Wijoyo, H., Putra, S. H., Watrianthos, R., Sinaga, R., Mardiana, R., ... & Indarto, S. L., "*Merdeka Kreatif di Era Pandemi Covid-19: Suatu Pengantar*", (Vol. 1). Green Press, 2020.
- [5] Nugraha, F. A., Harani, N. H., & Habibi, R., "*Analisis Sentimen Terhadap Pembatasan Sosial Menggunakan Deep Learning*", Kreatif, 2020.
- [6] Budiman, Dheni Apriansani., Nugraha, Decky Maulana., "*Aplikasi Raport Online Berbasis Web Menggunakan Framework CodeIgniter*", *Jurnal Computech & Bisnis STMIK Mardira Indonesia*, Bandung, 2019, (12 Juni 2021).
- [7] Elly Indrayuni, "*Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes*", *Jurnal Khatulistiwa Informatika*, Vol. Vii, No. 1 Juni 2019.
- [8] Taqwa Hariguna, Wiga Maulana Baihaqi and Aulia Nurwantia, "*Sentiment Analysis of Product Reviews as A Customer Recommendation Using the Naive Bayes Classifier Algorithm*", *International Journal of Informatics and Information Systems* Vol. 2, No. 2, September 2019.
- [9] Nurhikmah Nasution<sup>1</sup>, Ibnu Rasyid Munthe and Gomal Juni Yanris, "*Sentiment Analysis Of Detergen Products At Suzuya Mall Rantauprapat Navie Bayes Method*", *Jurnal Mantik*, 6 (1) (2022) 254-262, 2022.
- [10] Najma Sultana, Pintu Kumar, Monika Rani Patra, Sourabh Chandra and S.K. Safikul Alam, "*Sentiment Analysis For Product Review*", *Ictact Journal On Soft Computing*, April 2019, Volume: 09, Issue: 03, Doi: 10.21917/Ijsc.2019.0266.