

# *Sentiment Analysis about Product and Service Evaluation of PT Telekomunikasi Indonesia Tbk from Tweets Using TextBlob, Naive Bayes & K-NN Method*

Reza Hermansyah

Department of Technology Management  
Institut Teknologi Sepuluh Nopember  
Surabaya, Indonesia  
rezahrmsyh@gmail.com

Riyanarto Sarno

Department of Informatics  
Institut Teknologi Sepuluh Nopember  
Surabaya, Indonesia  
riyanarto@if.its.ac.id

**Abstract**—Online reviews are very important for any business that wants to control its online reputation. This allows businesses to have active and positive participation from consumers. As an information and communication company in Indonesia PT Telekomunikasi Indonesia Tbk commonly called Telkom require a customer's perspective or review to maintain the relevance of their digital products on the market. One method often used to analyze online reviews is sentiment analysis. Sentiment Analysis is used to gain an understanding of the opinions, attitudes, and emotions expressed in the mention of online by determining the emotional tone behind a series of words.

This research tries to compare classifications in sentiment analysis of Telkom's product from consumer reviews written in the form of tweets on Twitter. Each tweet about Telkom digital products such as Indihome, UseeTV, and Wifi.id will be collected as data. The use of classification types will be compared to help with the accuracy of sentiment analysis based on three types of methods TextBlob, Naïve Bayes & K-NN (K-Nearest Neighbor).

The best result of this research is the K-NN algorithm with an accuracy score of 75% followed by Naïve Bayes 69.44% and the last is TextBolb with 54.67%.

**Keywords** : *Online Review, Customer's Experience, Sentiment Analysis, TextBlob, Naïve Bayes, K-NN*

## I. INTRODUCTION

Intense competition in digital products in Indonesia requires every company that involved to keep improving its products. One way to improve the product by knowing reviews from consumers through social media. In this era, Twitter is one of the most frequently used microblogging and social networking service in expressing people's opinions.

Consumer satisfaction is important for both customers and businesses [1]. Customer satisfaction is used to measure the value of a brand to a product, service, or experience [2]. If

Telkom provides their customers with an astonishing customer service, they will stand out from the digital business competition in Indonesia. One method often used to analyze customer satisfaction is sentiment analysis [3]. Sentiment analysis is widely used because of its ability to gain rich insight into the details and the reason for otherwise opaque market trends.

Customer reviews written on the social media are type unstructured data, so they need appropriate techniques to be able to analyze them [4]. In this research, we choose TextBlob, Naïve Bayes & K-NN as classification methods because of ease the algorithm dan its implementation for sentiment analysis. Each tweet data will go through pre-processing such as punctuation removal, trimming, and stopword removal [5], [6] then classifying using TextBlob, Naïve Bayes & K-NN. The output of sentiment analysis in this research will be categories into two different emotions Positive or Negative [7].

This research showed how sentiment analysis of customer opinion determines customer satisfaction index and compares which classification the best used as a basis for corporate decision making for their products.

## II. LITERATURE REVIEW

Sentiment analysis is a type of data mining that deals with people's opinions through Natural Language Processing, computational linguistics, and text analysis [8]. In recent years, sentiment analysis is a great tool in social media to select up an outline of widespread public opinion at a precise subject. A lot of research has been done in the field of sentiment from hotel review [9] to movie review [10] which aims to extract opinions about topics, trends, etc.

Bayhaqy et al [11] used Decision Tree, K-Nearest Neighbor, and Naïve Bayes to classified tweets about E-Commerce. In their research K-NN algorithm gave the best result with data accuracy of around 88.50% compared to Decision Tree with 77% and Naïve Bayes approach with 64%.

Raut et al [12] classified sentiment from tweets using Optimized Feature Set to compare it with KNN and SVM. Their

experiment using 113 tweets dataset. The optimized feature gave the highest accuracy of about 82.00%.

Laksono et all [13] extracted customer reviews on TripAdvisor using Naïve Bayes and TextBlob to determine the sentiment of each comment. The result of the experiment shown Naïve Bayes have better data accuracy with 72.06% while TextBlob only obtained an accuracy of 69.12%.

Fikri et al [14] used SVM and SentiWordnet to classified on Google Playstore and Apple Appstore written in the Indonesian. Their experiment showed that the SVM algorithm achieves accuracies of 89% compared to SentiWordnet with only 56%.

Naïve Bayes is commonly used to calculate probability and statistical knowledge-based on Bayes theory [11], [15]. Meanwhile, K-NN knowns as an algorithm that calculates the distance between data testing to the nearest training data to classify each data test [16], [17].

The level of accuracy in this study will be determined by the classification method. As a comparison Texblob, Naïve Bayes, and K-NN are used to check the level of performance classification data. Variable such as True Positive, True Negative, False Positive, and False Negative as an element of confusion is the key to the succession of each method.

### III. PROPOSED METHOD

This research uses comments on twitter with the object is Telkom products such as Indihome, UseeTV, and Wifi.id as data sources. The data is processed by python programming language using Textblob, Naïve Bayes & K-NN Classification. Several steps need to be done to get the best results in analyzing sentiments. The steps consist of collecting, preprocessing, and classification sentiment analysis.

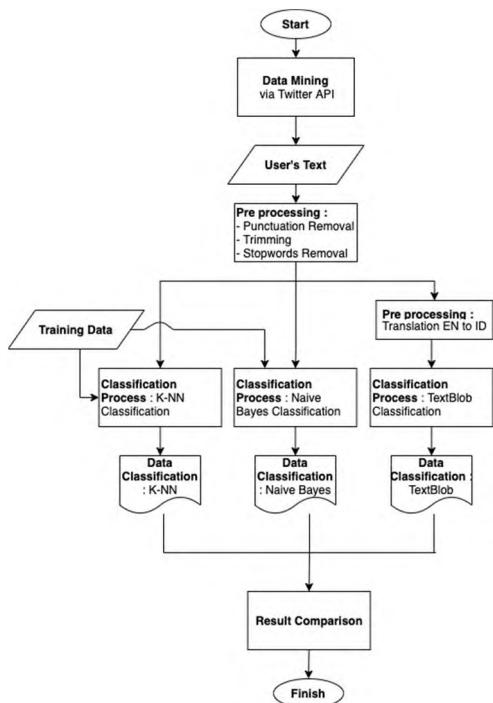


Fig. 1 Sentiment Analysis Process

#### A. Collecting Data

In this study, the data used for sentiment analysis was obtained from Twitter. User tweets will be collected in real-time based on product keywords with the help of python library called tweepy.

Python Installation :

\$ pip install tweepy

#### B. Labeling Data

After collecting the dataset, the next stage is giving labeling. The dataset will be divide into two classes and two parts. The classes determine the type of positive or negative sentiment. The parts of the dataset will be split into data testing and training. Training data is to adjust the parameters model and reduce the bias of predictions, while test data is to provide an unbiased evaluation of a final model. In this experiment we split the dataset into 100 data tests and 900 data training.

TABLE I. LABELING TABLE

Product	Tweet	Sentiment
Indihome & UseeTV	@indihome Untung ada UseeTV jadi engak boring	Positive
Wifi.id	Replying to @MahasiswaUMS Nevada kost. Wifi nya Wifi id telkom kenceng pol.	Positive
Indihome	@Indihome knpsi tiap malam bikin emosi	Negative
Indihome	@IndiHome gmna ini oi udh sebulan ngelag trus maen game.. parah sekali indihome..	Negative

#### C. Pre-Processing

Pre-processing is a stage where the data is prepared ready to be analyzed. The stages include punctuation removal, trimming, and stopword removal.

TABLE II . PRE-PROCESSING PROCESS

Pre-Processing	Information
Punctuation Removal	Process to removes all garbage item such as apostrophes, brackets, colons, commas, dashes, etc
Trimming	string manipulation to remove whitespace between words.
Stopword Removal	Removing a set of commonly used words in any language.
Translation EN to ID (for TextBlob)	Translate each word from English to Indonesian language

#### D. Classification

After the pre-processing process, the next stage of sentiment analysis is classification. This step is one of the keys to the accuracy of sentiment based on implement the use of training data to the algorithm.

- TextBlob

TextBlob is one of many python library for processing in the Natural Language Processing. Textblob can be used to perform a variety of NLP tasks ranging from parts of speech to sentiment analysis, and language translation to text classification [18].

Python Installation :

\$ pip install -U textblob

\$ python -m textblob.download\_corpora

TextBlob can calculate sentiment based on the polarity and subjectivity. It can determine whether it is a positive text or negative, if the polarity is  $> 0$  it is considered as positive while  $< 0$  is considered as negative.

A big advantage using TextBlob its provided a very easy interface and offers a lot of features like phase extraction, pos-tagging, sentiment analysis, etc.

One disadvantage of using this method is the absence of features such as dependency parsing, word vectors, etc.

- Naïve Bayes

Thomas Bayes formulated his famous probability theorem in the 18<sup>th</sup> century called Bayes theorem. Bayes theorem based on prior knowledge of conditions that might be associated with the event to describes the probability of an event. This type of classifier assumes that the presence of a particular feature of a class is unrelated to the presence of any other feature [19]. An explanation of how Naive Bayes classifier works can be seen in the formula (1).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

where :

- $P(A|B)$  = the posterior probability
- $P(B|A)$  = the likelihood probability
- $P(A)$  = the prior probability
- $P(B)$  = evidence

Advantage of Naïve Bayes as classifier:

- Requires a small amount of training data to estimate test data.
- Perform better when the assumption of independent predictors holds true.

- K-NN (K-Nearest Neighbors)

K-NN is a classification algorithm for a collection of data based on the prior classification of data learning. K-NN is stated mathematically as the following relationship (2). This method will compare the data test with the distance between data testing to the closest pattern space from the training data [5].

$$d(x, y) = \left( \sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (2)$$

where :

$p = 1$ , Manhattan Distance

$p = 2$ , Euclidian Distance

$p = \infty$ , Cybychev Distance

$x_i$  = datapoint from dataset

$y_i$  = new data point (to be predicted)

Advantage of K-NN as classifier:

- No training period, K-NN stores the training dataset and only learn when need to make real-time prediction.
- New data can be added seamlessly because the K-NN algorithm requires no training before making predictions.

#### E. Evaluation

The efficiency of sentiment analysis submission is calculated through different classification experiments in form of tweet data. Parameters like accuracy, precision, and recall will be performed to determine which classification the best for this study [20].

Where Naïve Bayes and K-NN classified as supervised machine learning we can calculate results with those three parameters based on training data. While Textblob classified as Natural Language Processing can work without training data.

In the field of machine learning, a confusion matrix is used to describe the performance of a classification model. Confusion matrix classified data into four variable such as True Positive, True Negative, False Positive and False Negative, the number of correct and incorrect predictions are summarized with count values and broken down by each class.

TABLE III . CONFUSION MATRIX

		Predicted	
		Positive Review	Negative Review
Actual	Positive Review	$TP^a$	$FN^b$
	Negative Review	$FP^c$	$TN^d$

a. TP (True Positive) : Result of positive data that is predicted to be correct

b. FN (False Negative) : Result of positive data but predicted as negative data

c. FP (False Positive) : Result of negative data but predicted as positive data

d. TN (True Negative) : Result of negative data that is predicted to be correct

- Accuracy

The accuracy of measurement in statistics is how close a result comes to the true value. To calculate the accuracy value can use the equation :

$$A = \frac{(TP + TN)}{(TP + FP + TN + FN)} \times 100\% \quad (3)$$

- Precision

Precision is the portion closeness of a set of answers to others presented by the experimenter. In mathematical calculations, precision can be formulated from the quotient of True Positive (TP) and total of True Positive and False Positive (FP).

$$P = \frac{TP}{(TP + FP)} \times 100\% \quad (4)$$

- Recall

While recall or sensitivity is defined as true predicted data to all existing data that are positive (including False Negative). The recall formula is :

$$R = \frac{TP}{(TP + FN)} \times 100\% \quad (5)$$

#### IV. RESULT

This section describes our experimental result and analyzes the performance of each classification method.

In this study, we use 900 tweets as a training dataset. The Textblob method which uses natural language processing measures all of the training dataset into a confusion matrix. While Naïve Bayes and K-NN classify as Machine learning divide into 80:20 for training and data test to confusion matrix.

TABLE IV . CONFUSION MATRIX FOR TEXTBLOB

		Predicted	
		Class	Positive
Actual	Positive	64	404
	Negative	4	428

TABLE V . CONFUSION MATRIX FOR NAIVE BAYES

		Predicted	
		Class	Positive
Actual	Positive	49	38
	Negative	7	86

TABLE VI. CONFUSION MATRIX FOR K-NN

		Predicted	
		Class	Positive
Actual	Positive	72	15
	Negative	40	53

With the help of confusion matrices, parameter accuracy, precision and recall can be calculated using formulas (3),(4),(5).

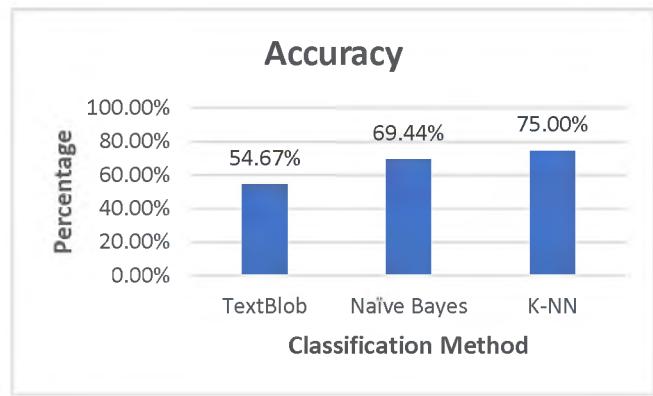


Fig. 2 Accuracy comparison of each method

Shown from figure 2 TextBlob is the less accurate classification from others. The basis of TextBlob which calculates parts per word in the English language have difficulty interpreting slang variations in Indonesian tweets. While Naïve Bayes and K-NN methods that have a machine learning basis can compare each data test (tweets) with training data provided in the Indonesian language.

K-NN method performs better with accuracy 75% because of its ability to find similarity between observations and also its inherent nature to optimize locally.

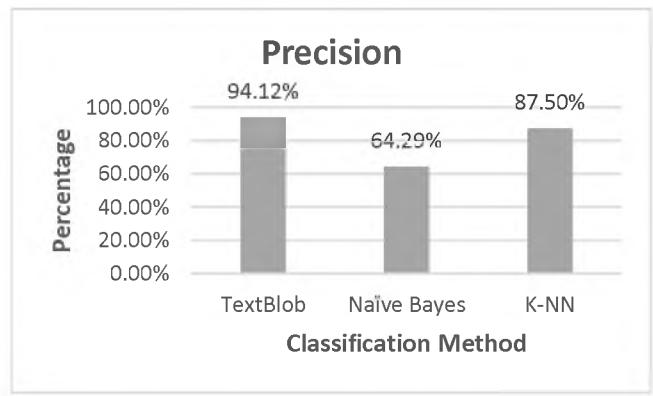


Fig. 3 Precision comparison of each method

Figure 3 shows that TextBlob has better precision rather than other methods, this happens because of NLP basis forces the TextBlob to calculate the entire dataset as a data test. While the result of precision from machine learning-based shows K-NN performs better than Naïve Bayes with 87.50% compared to 64.29%. K-NN classifier has more confidence in the portion closeness to the real answer.

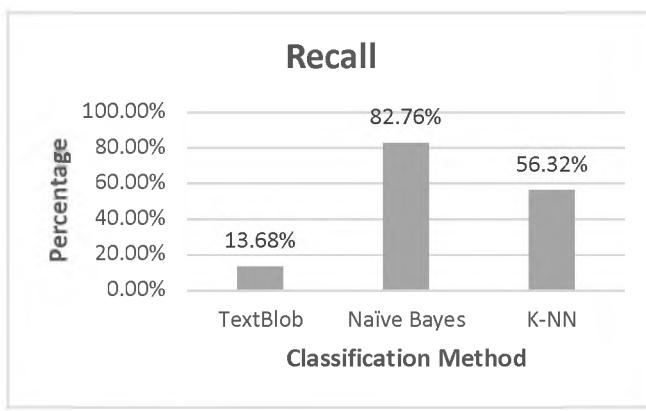


Fig. 4 Recall comparison of each method

Figure 4 shows that TextBlob has inversely proportional to precision in figure 3. The NLP causes TextBlob to calculate the entire dataset as a data test, which makes the lowest recall compare to other methods. Recall of machine learning-based shows that Naïve Bayes is superior to K-NN, its defined Naïve Bayes calculate better true predicted data to all existing data that mark as true positive.

Overall it can be seen that the comparative value of each category of tweet respect to each classification method. The accuracy of TextBlob, Naïve Bayes, and K-NN by 54.67%, 69.44%, and 75%. Precision of TextBlob, Naïve Bayes and K-NN by 94.12%, 64.29%, and 87.50%. While recall result of TextBlob, Naïve Bayes and K-NN by 13.68%, 82.76%, and 56.32%. From the result shown K-NN has better classification with higher accuracy than other methods.

The next scenario is to evaluate Telkom product using K-NN which performing better as classification than other methods. Indihome, UseeTV and Wifi.id are used as keywords to crawling 100 data tweets for each product from Twitter.

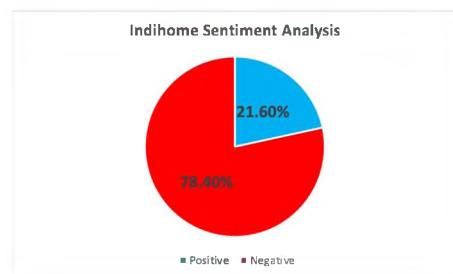


Fig. 5 Indihome Sentiment Analysis using K-NN Classifier

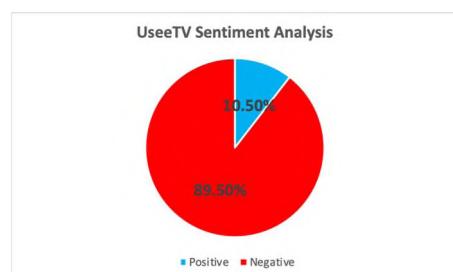


Fig. 6 UseeTV Sentiment Analysis using K-NN Classifier

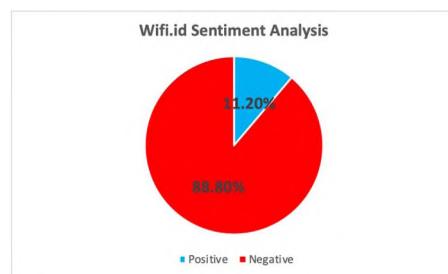


Fig. 7 Wifi.id Sentiment Analysis using K-NN Classifier

Figures 5,6 & 7 shown sentiment charts about Telkom products using K-NN as a classifier. The result of the Indihome product has 78.40% negative sentiment following with 21.60% positive sentiment. UseeTV also dominated by negative sentiment with 89.50% and 10.50% for positive sentiment. The last product is Wifi.id, it has 88.80% negative sentiment and 11.20% for positive sentiment.

## V. CONCLUSION

In this study, we create a sentiment classifier from twitter on various customer opinions about Telkom products. Three different types of classifiers TextBlob, Naïve Bayes, and K-NN have been processed to get the best result since consumer satisfaction is important for both customers and businesses.

From the experiment, shown that K-NN classification has the better accuracy than others. The result show the accuracy of the Textblob, Naïve Bayes and K-NN amounted to 54.67%, 69.44%, and 75%. The Textblob method performs worse than Naïve Bayes and K-NN in the recall process due to a non-supervised type classifier compared to a machine learning-based classifier by 13.68%, 82.76%, and 56.32%. However in the recall process give advantage for Textblob force it to calculate the entire dataset as a data test give the precision result of Textblob, Naïve Bayes, and K-NN by 94.12%, 64.29%, and 87.50%. As we compare each classifier K-NN method gives more accurate prediction, so we can conclude that K-NN is performing better than Naïve Bayes and Textblob. K-NN method performs better because of its ability to find similarity between observations and its inherent nature to optimize locally.

The experiment hypothesis showed negative sentiment dominates the customer opinion through twitter to Telkom product. The result showed negative sentiment for Indihome, UseeTV, and Wifi.id using K-NN as classifier are 78.40%, 89.50%, and 88.80%. Based on customer tweet data, negative sentiment arises from complaints about poor product performance and a lack of product features.

From the experiment suggested that Telkom need to focus on daily customer feedback to increase customer loyalty to their products. It's important for Telkom to Improving and keeping a positive attitude toward critics about their product to gain back customer trust.

Further improvement can be done by using a larger and more complex train or review dataset to increase the value of accuracy

in sentiment analysis. Future studies also can include each aspect of the product to improve the sentiment hypothesis.

## REFERENCES

- [1] A. Ghazian, M. H. Hossaini, and H. Farsijani, "The Effect of Customer Relationship Management and its Significant Relationship by Customers' Reactions in LG Company," *Procedia Econ. Finance*, vol. 36, pp. 42–50, 2016, doi: 10.1016/S2212-5671(16)30014-4.
- [2] I. Buil, E. Martínez, and L. de Chernatony, "The influence of brand equity on consumer responses," *J. Consum. Mark.*, vol. 30, no. 1, pp. 62–74, Jan. 2013, doi: 10.1108/07363761311290849.
- [3] B. Pang and L. Lee, "Opinion mining and sentiment analysis," p. 94.
- [4] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *Int. J. Inf. Manag.*, vol. 35, no. 2, pp. 137–144, Apr. 2015, doi: 10.1016/j.ijinfomgt.2014.10.007.
- [5] A. R. Naradhipa and A. Purwarianti, "Sentiment Classification for Indonesian Message in Social Media," p. 4.
- [6] D. A. K. Khotimah and R. Sarno, "Sentiment Detection of Comment Titles in Booking.com Using Probabilistic Latent Semantic Analysis," in *2018 6th International Conference on Information and Communication Technology (ICoICT)*, Bandung, May 2018, pp. 514–519, doi: 10.1109/ICoICT.2018.8528784.
- [7] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A Library for Large Linear Classification," p. 31.
- [8] V. Bonta, N. Kumaresh, and N. Janardhan, "A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis," p. 6, 2019.
- [9] R. Priyatina and R. Sarno, "Sentiment Analysis of Hotel Reviews Using Latent Dirichlet Allocation, Semantic Similarity and LSTM," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 4, pp. 142–155, Aug. 2019, doi: 10.22266/ijies2019.0831.14.
- [10] Suhariyanto, A. Firmanto, and R. Sarno, "Prediction of Movie Sentiment Based on Reviews and Score on Rotten Tomatoes Using SentiWordnet," in *2018 International Seminar on Application for Technology of Information and Communication*, Semarang, Sep. 2018, pp. 202–206, doi: 10.1109/ISEMANTIC.2018.8549704.
- [11] A. Bayhaqy, S. Sfenrianto, K. Nainggolan, and E. R. Kaburuan, "Sentiment Analysis about E-Commerce from Tweets Using Decision Tree, K-Nearest Neighbor, and Naïve Bayes," in *2018 International Conference on Orange Technologies (ICOT)*, Nusa Dua, BALI, Indonesia, Oct. 2018, pp. 1–6, doi: 10.1109/ICOT.2018.8705796.
- [12] R. K. P. Anup Raut, "Sentiment Analysis using Optimized Feature Sets in Different Twitter Dataset Domains," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 11, pp. 3035–3039, Sep. 2019, doi: 10.35940/ijitee.K2195.0981119.
- [13] R. A. Laksono, K. R. Sungkono, R. Sarno, and C. S. Wahyuni, "Sentiment Analysis of Restaurant Customer Reviews on TripAdvisor using Naïve Bayes," in *2019 12th International Conference on Information & Communication Technology and System (ICTS)*, Surabaya, Indonesia, Jul. 2019, pp. 49–54, doi: 10.1109/ICTS.2019.8850982.
- [14] M. Fikri and R. Sarno, "A comparative study of sentiment analysis using SVM and SentiWordNet," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 13, no. 3, p. 902, Mar. 2019, doi: 10.11591/ijeecs.v13.i3.pp902-909.
- [15] H. Zhang, Z.-X. Cao, M. Li, Y.-Z. Li, and C. Peng, "Novel naïve Bayes classification models for predicting the carcinogenicity of chemicals," *Food Chem. Toxicol.*, vol. 97, pp. 141–149, Nov. 2016, doi: 10.1016/j.fct.2016.09.005.
- [16] M. Jupri and R. Sarno, "Taxpayer compliance classification using C4.5, SVM, KNN, Naïve Bayes and MLP," in *2018 International Conference on Information and Communications Technology (ICOIACT)*, Yogyakarta, Mar. 2018, pp. 297–303, doi: 10.1109/ICOIACT.2018.8350710.
- [17] B. Y. Pratama and R. Sarno, "Personality classification based on Twitter text using Naïve Bayes, KNN and SVM," in *2015 International Conference on Data and Software Engineering (ICoDSE)*, Yogyakarta, Indonesia, Nov. 2015, pp. 170–174, doi: 10.1109/ICoDSE.2015.7436992.
- [18] V. S and J. R, "Text Mining: open Source Tokenization Tools – An Analysis," *Adv. Comput. Intell. Int. J. ACII*, vol. 3, no. 1, pp. 37–47, Jan. 2016, doi: 10.5121/acii.2016.3104.
- [19] P. Tripathi, S. Kr. Vishwakarma, and A. Lala, "Sentiment Analysis of English Tweets Using Rapid Miner," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, Jabalpur, India, Dec. 2015, pp. 668–672, doi: 10.1109/CICN.2015.137.
- [20] Y. Jiao and P. Du, "Performance measures in evaluating machine learning based bioinformatics predictors for classifications," *Quant. Biol.*, vol. 4, no. 4, pp. 320–330, Dec. 2016, doi: 10.1007/s40484-016-0081-2.