Wella Wella

Is the Internet in Indonesia Has a Good Sentiment from Netizen?

Quick Submit

Quick Submit

Universitas Multimedia Nusantara

Document Details

Submission ID
trn:oid:::1:3430340673

Submission Date
Dec 2, 2025, 11:21 AM GMT+7

Download Date
Dec 2, 2025, 11:29 AM GMT+7

File Name
prosiding_Is_the_Internet_in_Indonesia_Has_a_Good.pdf

File Size
389.5 KB

6 Pages

3,879 Words

20,703 Characters

13% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Cited Text
- Small Matches (less than 8 words)

Exclusions

- 2 Excluded Sources

Match Groups

36 Not Cited or Quoted 13%
Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

7% Internet sources
11% Publications
5% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **36** Not Cited or Quoted 13%
  Matches with neither in-text citation nor quotation marks
- **0** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

7% Internet sources
11% Publications
5% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Publication	
Subasish Das. "Artificial Intelligence in Highway Safety", CRC Press, 2022		3%
2	Student papers	
Universitas Brawijaya		1%
3	Internet	
link.springer.com		<1%
4	Internet	
www.sciencegate.app		<1%
5	Publication	
Muhammad Hasnain, Muhammad F. Pasha, Imran Ghani, Muhammad Imran, Mo...		<1%
6	Publication	
Aria Eka Putra, Suryasari. "The Design of Sentiment Analysis Application using To...		<1%
7	Publication	
Lincoln M. Barcelo, Cherry Mae P. Bairan, Rhovie Ann M. Semillano, Ronn-Tristan ...		<1%
8	Internet	
ejournals.umn.ac.id		<1%
9	Student papers	
King Mongkut's University of Technology Thonburi		<1%
10	Student papers	
University of Lancaster		<1%

11 Internet
library.oapen.org <1%

12 Publication
Ashima Yadav, Ayush Agarwal, Dinesh Kumar Vishwakarma. "XRA-Net Framewor... <1%

13 Publication
Fazjar Sekti Aji, Ade Iriani, Hendry. "Toxicity and Sentiment Analysis About Digita... <1%

14 Publication
Praise Setiawan Saragih, Deden Witarsyah, Faqih Hamami, Jose Manuel Machado.... <1%

15 Internet
www.ijosi.org <1%

16 Publication
Rizkiansyah, Asti Herliana, Doni Purnama Alamsyah, Tjia Fie Tjoe. "Comparison of... <1%

17 Internet
archive.org <1%

18 Internet
biomedres.us <1%

19 Internet
journal.uinjkt.ac.id <1%

20 Internet
www.coursehero.com <1%

21 Publication
Evelyn Sierra, Dini Adni Navastara, Diana Purwitasari. "A Combination of Term Fr... <1%

22 Publication
Mikogizka Satria Kartika, P. H. Gunawan, Indwiarti, Narita Aquarini. "Detection o... <1%

23 Publication
Sulthan Rafif, Pramana Yoga Saputra, Moch Zawaruddin Abdullah. "Classification ... <1%

24 Publication
Titus Yory Datubakka, Istikmal, Arif Indra Irawan. "Comparison Analysis Of K-Ne... <1%

25 Publication

Zhiliang Zhu, Jie Liang, Deyang Li, Hai Yu, Guoqi Liu. "Hot Topic Detection Based o... <1%

26 Internet

d-nb.info <1%

27 Internet

ejurnal.stmik-budidarma.ac.id <1%

28 Internet

www.hindawi.com <1%

29 Internet

www.warse.org <1%

Is the Internet in Indonesia Has a Good Sentiment from Netizen?

Kornelius[1], Wella[2]
[1,2] Information System Departement, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia
[1] kornelius@student.umn.ac.id
[2] wella@umn.ac.id

***Abstract*— During this pandemic, the Internet has become a basic necessity that equals clothing, food, and shelter. People use the Internet for various activities such as ordering food, communicating, learning remotely (online), working, and social media, one of which is Instagram. On Instagram, people can upload pictures, videos, and even text if they already have an account. This feature is not only limited to an individual but can also be used by companies or, in this case, internet providers. On Instagram comments on internet provider accounts, there are often positive and negative comments that overlap in the comments section, making it difficult for providers to measure the satisfaction of their customers. Therefore in this report, researchers will use sentiment analysis with the Naïve Bayes algorithm and Support Vector Machine (SVM) to measure customer satisfaction through the many hate speeches in the Instagram comments section of internet provider accounts. The research will begin by collecting 1000 *data* on the internet provider Instagram comments (Text Mining), then the existing *data* will enter the sentiment analysis process. After that, the sentiment results will be inputted into two algorithms, namely Naïve Bayes and Support Vector Machine (SVM). Then the accuracy will be calculated using the K-Fold Cross Validation and Confusion Matrix to become a precise result. The research results will be in the form of a percentage of customer satisfaction and dissatisfaction based on the many hate speeches that each internet provider has. The results of the analysis of this study are expected to help internet providers in Indonesia know that many customers who are not satisfied with their services can improve the quality of their services.**

***Index Terms*—Provider, Internet, Instagram, SVM, Naïve Bayes, Sentiment Analysis**

I. INTRODUCTION

The Internet has become a primary need in this era. In Indonesia, the Internet has been used by more than half of the Indonesian people. In January 2020, it was estimated that approximately 175.4 million, or 64% of the total population of 272.1 million people, have accessed the Internet [1]. The Internet can be used freely, not only for positive things but also for negative things such as expressing hate speech. According to Teguh Afriyadi, Acting Director of Informatics Application Control at the Ministry of Communication and Informatics, the Internet has become a popular platform for people, in general, to spit hate speech because the laws applied are still not accommodating [2]. This usually happens most often on Social Media.

It is recorded that 160 million Indonesians actively use Social Media, and on average, they use or open Social Media for 3 hours, 26 Minutes a day[1]. People use various kinds of social media in their daily life, such as Youtube (88%), Whatsapp (84%), Facebook (82%), Instagram (79%), and Twitter (56%), where they can socialize and engage in social activities. Cyberspace. Internet speed affects the smooth running of activities on Social Media, but Indonesia's Internet speed is among the two lowest out of 45 countries with an average speed of 15.5 Mbps [3]. Not only is it not fast enough, but there are other problems, such as connection loss and unanswered complaints from customers, that make them feel less satisfied with the service providers they have paid for [4]. Therefore, it can be concluded that the performance of service providers can be determined by the many complaints or hate speech they get from their respective provider customers.

Comments given by customers can usually be found on certain social media, one of which is Instagram. In this study, the comments on the internet provider account post will determine how satisfied the provider's service is. Each internet provider's Instagram account receives comments every day on their posts with many negative and positive comments. Therefore, sentiment analysis is needed to process large amounts of data. In this study, the algorithm used is Naive Bayes and Support Vector Machine (SVM). Naive Bayes is used to classifying with assumed independent variables and produce many combinations[5]. At the same time, SVM is a classification that aims to separate two classes using a hyperplane-linear classifier[6]. With these two algorithms, the comments obtained can be separated so that it is known which comments are classified as Hate speech or not and are ready to calculate their accuracy.

II. Basic Theory

A. Text Mining

Text mining is an intensive process where users can interact with documents and use analytical tools to find patterns of interest. In-Text Mining, researchers usually approach finding knowledge in large amounts of text [7]. The text mining process consists of several steps [7]:

• Defining the problem

At this stage, the problem in the domain needs to be understood, and the existing questions need to be answered or defined.

• Collecting the necessary data

The data collected must be as desired and identified. Data also needs to be documented so that it is ready for further analysis.

• Defining features

The definition of text that has characterized the function of the text itself.

• Analyzing the data

The stage where the data is processed so that existing patterns can be found. To solve a problem, usually, several models are needed. Every model has different characteristics. The model used can be a white box and a black box. A suitable model will usually depend strongly on the data.

• Interpreting the results

The results can be found in the analysis. At this stage, verification and validation are needed to improve the reliability of the results.

B. Sentiment Analysis

Sentiment Analysis in Natural Language Processing (NLP) builds a system to recognize and extract opinions in text form. Information in the form of text is currently widely available on the Internet in the format of forums, blogs, social media, and sites containing reviews. With the help of sentiment analysis, previously unstructured information can be transformed into more structured data.

The data can explain people's opinions about products, brands, services, politics, or other topics. Companies, governments, and other fields then use this data to make marketing analyses, product reviews, product feedback, and public services [8].

C. Naïve Bayes Classifiers

According to the theory, the optimal Bayesian classifier gives the best classification results based on the given hypothesis. This is mainly (but not only) true in text mining because assigning rect class-label c j depends on the number of attributes, i.e., words in social media contributions or text documents. After collecting a large number of training samples (which is good from the point of computing probabilities vie), their vocabulary may be huge - thousands or tens of thousands of unique words (or generic terms). Each word has its probabilities in each class, and there are many possible combinations of attributes, even if the data is reduced by, for example, eliminating unimportant syllables.

For large values of N, the computational complexity is given by the above items. Due to the possible conditional interdependence of the attributes, it is necessary to calculate the probabilities of all their possible combinations for a given cj. An idea called a nave Bayes classifier, which is (theoretically not accurate enough) based on the assumption that there are no interdependencies between attributes (i.e., 'nativity), was generated to make computing more accessible and practically applicable. For classification, it is possible to use a more straightforward equation to determine cMAP, referred to as Naïve Bayes [7].

D. Support Vector Machine

Support Vector Machines, SVM, is another successful algorithm that has proven and demonstrated outstanding results in many different real-world applications. Its foundation dates back to relatively early times, beginning in 1963 when Russian (Soviet) mathematician Vladimir Vapnik (along with Alexey Chervonenkis, Alexander Lerner, and others) published statistical learning and pattern recognition work. These works gradually gave rise to one of the currently successful algorithms based on linear algebra theory. In principle, SVM is a classification algorithm that aims to separate two classes using a hyperplane - a linear classifier. There are several ways to place a hyperplane (or straight line in two-dimensional space) that separates elements of the two classes. If such an option exists, one can generally create many such constraints for a given set of training examples. First, according to the SVM idea, finding the most comprehensive zone (belt) separating the two classes is necessary. Properly selected examples called support vectors to define the edges of this zone [7].

E. Rapid Miner

It is an application with functions for machine learning, data preparation, text mining, and predictive analysis. These applications are commonly used for different purposes such as training, education, business and commercial, and others [10].

F. Data Pre-processing

This step must be done so that the data can be ready to be processed. At this stage, there are usually 5 Sub-Processes, namely Data cleansing, which makes irrelevant data relevant, Case folding makes letters into the same shape, tokenization separates sentences into tokens, stop words removal, and finally stemming where the process removes suffixes and affixes [9].

G. *K-Fold Cross-Validation*

10-fold cross-validation is the standard for incrementally testing the accuracy of classifiers. There are ten steps, each using 10% of the sample for testing and 90% for training, and each sample gradually takes part in both parts. Classification performance metrics are calculated at each step. The average of all ten measurements gives the final score. In our example, only the accuracy values are considered for

simplicity. The mean standard error (SEM) used in statistics to express the mean error for a test sample selected randomly from the data population (generalized error for the entire population) was also calculated. Given below is an example of R-code. A reader will notice that the code for cross-validation testing is similar to code that uses test sets for testing [7].

H. *Confusion Matrix*

Confusion Matrix is a computational method that is often used in machine learning to classify the behavior of classification models. The confusion matrix is usually applied to evaluate the performance of classifications on datasets—confusion Matrix. Some researchers use the Confusion Matrix to distinguish the predicted value and the original value of the model elements in software engineering. In the Confusion Matrix, there are 4 measurement methods, namely TP, FP, TN, and FN, used in Java programs [11].

I. *TF-IDF*

TF-IDF is an algorithm that is widely used for feature selection in text information processing. TF stands for Term Frequency which represents how often feature terms appear from a collection of texts. IDF stands for Inverse Document Frequency, which measures the general importance of a term that often appears in the data set [12].

J. *Instagram*

One of the social media that is often used by the public is Instagram, where users can share photos and videos for free with other Instagram users. Users can also view, comment, and like photos provided by other users. In this study, comments on Instagram will be used to retrieve data using the Instagram API. With this API, it is possible to retrieve comment data on Instagram [13].

K. *Tableau*

Tableau is an application that is used to present data with a more attractive appearance. Tableau users can use data sources virtually and process data in various forms, sizes, and types [14].

L. *Pycharm*

Pycharm is a popular application that is commonly used to write Python scripts. This application is also possible to import libraries such as pandas, numpy, sklearn, tinker, mtlab plot. Users only need to go through the code created to see the code results [15].

III. Research Methodology

A. *Research Object*

The objects studied are comments from Instagram posts from Indonesian internet provider accounts such as Indihome, Firstmedia, XLhome, Biznet, and Oxygen.id. From the results of these comments, you will see how much hate speech is posted from the Instagram account of their provider's product users. A lot of "hate speech" data will determine how much customer satisfaction from each provider.

- IndiHome (@indihomecare)
- FirstMedia (@firstmediaworld)
- XL Home (@xlhomeid)
- Biznet (@biznethome)
- Oxygen.id (@oxygen.id)

B. *Collecting Data*

Data collection will be carried out using PyCharm as an IDE to perform text mining using Open Source Scripts. The data to be retrieved will be obtained through the comment section on the Instagram post of the Indonesian internet provider account. Data that is considered valid is data in the post of the Indonesian internet provider's Instagram account.

C. *Data Preprocessing*

In this process, the data will be separated based on the type of data.

- Remove Duplicate Data: At this stage, the data will be checked whether there are similarities or repeated with each other and eliminated if there are similarities.
- Labeling Data: The data obtained will be labeled "hatespeech" on text containing hate words and "non-hate speech" on words that do not contain hate.
- Case Folding: All letters from the commentary sentences obtained are added to the lower case so that all letterforms are the same and can be read.
- Filtering: Reducing words or attributes that are less influential in the searched sentence to reduce the time when checking. As an example:
  - Usernames / Tags
  - Url / Link

- Stopwords Removal: Eliminate irrelevant words such as "Lol" or "Wkwkwk" to avoid interfering with the following process.

D. SVM/Naïve Bayes Modeling

At this stage, we will use the train data set obtained from GitHub so that it is ready to be modeled using SVM and Naive Bayes.

E. Calculation Accuracy

At this stage, the accuracy of the two algorithms will be calculated as a benchmark for the success of the algorithm model on the data being tested. The method used is K-Fold Cross-Validation, where the higher the percentage obtained, the better the results obtained.

F. Implementation of SVM/Naïve Bayes Model to Unlabelled data

After cross-validation on the train data set, the SVM/Naïve Bayes algorithm model is ready to be implemented on preprocessed data. The data will produce a percentage of accuracy and excel with a list of comments that can be considered hate speech and not.

G. Visualization Data

The excel data obtained will be processed or visualized to make it more straightforward to obtain many negative comments. The visualization will use Tableau.

IV. Result & Discussion

A. Problem Identification

This first stage is carried out to determine the research topic by analyzing related previous journals. Based on the journals collected and analyzed, it was found that Sentiment Analysis is often used to analyze people's responses to an object. Public response can be easily obtained through various media, for example, Instagram and Twitter. On Twitter and Instagram, there are many kinds of responses from the public, including hate speech. Based on previous journals, it was also found that Twitter responses were used more for sentiment analysis research than Instagram.

B. Determining Research Objectives

The second stage will provide solutions to the problems that have been described in the problem identification stage. Therefore, in this study, sentiment analysis will be carried out on hate speech given by the public to the Instagram account of internet providers in Indonesia. Instagram is used to test whether social media other than Twitter can be used in sentiment analysis or not and if the results obtained are valid and can be used. The data will be feasible to be presented.

C. Determining Research Objectives

This study aims to determine the feasibility of sentiment analysis results on Instagram comments on Indonesian internet provider accounts. The feasibility will be determined from the accuracy, precision, and recall values obtained from the confusion matrix. In addition, this study will compare two algorithms with a relatively high level of accuracy.

D. Pre-processing Data

At the data collection stage, data collection will be carried out using the pycharm application. Pycharm is run with source code specially made to pull data from Instagram. The source code results will show the user's username and the text containing the user's comments in the comment section.

The data obtained is still dirty and has data that is not used for research, such as user id, comment id, profile URL, profile pic URL, and date. After that, the results obtained will be cleaned at the Cleansing Data stage.

Duplicate data and missing attributes will be removed to become more valid by using the rapidminer application tools. Several filters will be carried out at this stage to remove URLs such as website links, Usernames from comments such as '@,' and emoticons contained in one of the text data. This stage is done to reduce the less important characters so that the data will become more valid.

At this stage, first case folding will be carried out where all letters in the text will be made lowercase. Then the second tokenization will be carried out where, when executed, it will separate sentences into fractional words, for example, "It's been three days my internet is dead" will change to "already," "three," "days," "internet," "me," and " die." The next stage is stopword removal, where its function includes removing words or letters that are not important. For example, in the word "same pny, I also have a bill that is not appropriate" there are "pny" which means "have," "sy" which means "I," "jg" which means "too," "g" which means "not" will be changed to be "same bill match." The last stage in preprocessing is stemming, where the words will be converted into their basic form. For example, the words "Min, please dm for a new installation, etc. Thank you" will become "Min please dm for a new install, etc. Thank you," and the data has been cleaned and is ready to enter the data processing stage.

E. Processing Data

After the data has been prepared, the data will enter the separation stage using K-Fold Cross-Validation, where 10% of the data will be used as testing data. The remaining 90% will be used as

training data, and the process will continue to be repeated up to 10 times with a change of 10 % new testing data every iteration. The clean data that has been obtained will be entered and then processed using K-Fold Cross-Validation. Cross-Validation can be set for the use of the algorithm you want to use. In this process, we will use Naive Bayes and SVM and generate a model of each algorithm. At this stage, the first will use Naive Bayes as an algorithm for searching the first model.

The data will be divided into training and testing. The Naive Bayes model will be applied in the testing section, and the performance will be tested on the data.

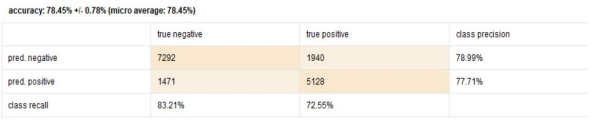

Fig. 1. Naïve Bayes Model Result

As shown in Figure 1, the accuracy is 78.45%, precision is 77.72%, and recall is 72.55%. After that, the same process will be repeated but using SVM as the algorithm used.

The SVM model will be applied in the testing section, and the performance will be tested against the data.

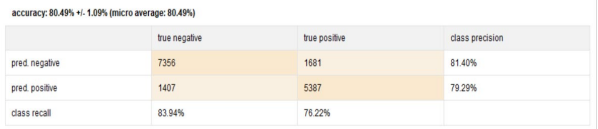

Fig. 2. SVM Model Result

As shown in Figure 2, the accuracy is 80.49%, precision is 79.29%, and recall is 76.22%. At this stage, data classification will be carried out. The model created from Naive Bayes and SVM will classify whether the data given is positive or negative.

At this stage, data classification will be carried out. The model created from Naive Bayes and SVM will classify whether the data given is positive or negative. First, the data will be classified using Naive Bayes and SVM to get results like this:

TABLE I. PROVIDER NAÏVE BAYES RESULT

Provider	*Total*	Naïve Bayes				
		Positive	*Percentage*	*Negative*	*Percentage*	Interval
Indihome	4116	926	22%	3190	78%	55%
Firstmedia	1184	582	49%	602	51%	2%
Xlhome	1055	338	32%	717	68%	36%
Biznet	1522	679	45%	843	55%	11%
Oxygen.id	1025	378	37%	647	63%	26%

TABLE II. PROVIDER SVM RESULT

Provider	*Total*	SVM				
		Positive	*Percentage*	*Negative*	*Percentage*	Interval
Indihome	4116	1335	32%	2781	68%	35%
Firstmedia	1184	526	44%	658	56%	11%
Xlhome	1055	426	40%	629	60%	19%
Biznet	1522	725	48%	797	52%	5%
Oxygen.id	1025	469	46%	556	54%	8%

With the results obtained,

- Indihome has 926 positive comments and 3190 negative comments using Naive Bayes. Then Indihome has 1335 positive comments and 2781 negative comments using SVM.
- First media has 582 positive comments and 602 negative comments using Naive Bayes. Then Firstmedia has 526 positive comments and 658 negative comments using SVM.
- Biznet has 679 positive comments and 843 negative comments using Naive Bayes. Then Biznet has 725 positive comments and 797 negative comments using SVM.
- XLhome has 338 positive comments and 717 negative comments using Naive Bayes. Then XLhome has 426 positive comments and 629 negative comments using SVM.
- Oxygen.id has 378 positive comments and 647 negative comments using Naive Bayes. Then Oxygen.id has 469 positive comments and 556 negative comments using SVM.

V. CONCLUSION

Based on all the current results, it can be concluded that SVM is a more suitable algorithm for this study with 80.49% accuracy rather than Naive Bayes with 78.45% accuracy. It can be said that Indonesian provider needs to improve their services because the negative comment is more dominant than positive comment based on the result.

The results of this study can be used as a basis for further research by adding the amount of data in the next year to improve the quality of the results obtained. It is also possible to compare the results with other classification algorithms such as lexicon.

REFERENCES

[1] S. Kemp, "DIGITAL 2021: INDONESIA," 2021. https://datareportal.com/reports/digital-2021-indonesia (accessed Nov. 30, 2020).

[2] R. Puspita, "Kemenkominfo: Internet Paling Populer Sebarkan Hate Speech," 2020. https://republika.co.id/berita/qi77o7428/kemenkominfo-internet-paling-populer-sebarkan-hate-speech (accessed Nov. 30, 2020).

[3] CNN Indonesia, "Kecepatan Internet Indonesia ke-2 Terbawah dari 45 Negara," 2020. https://www.cnnindonesia.com/teknologi/20190201152441-185-365734/kecepatan-internet-indonesia-ke-2-terbawah-dari-45-negara (accessed Nov. 30, 2020).

[4] CNN Indonesia, "Netizen Cibir Jaringan IndiHome Sering Lambat Kala WFH Corona," 2020.

https://www.cnnindonesia.com/teknologi/20200409181907-192-492233/netizen-cibir-jaringan-indihome-sering-lambat-kala-wfh-corona (accessed Nov. 30, 2020).

[5] H. Sudira, A. L. Diar, and Y. Ruldeviyani, "Instagram Sentiment Analysis with Naive Bayes and KNN: Exploring Customer Satisfaction of Digital Payment Services in Indonesia," 2019 Int. Work. Big Data Inf. Secure. IWBIS 2019, pp. 21–26, 2019, doi: 10.1109/IWBIS.2019.8935700.

[6] S. Naz, A. Sharan, and N. Malik, "Sentiment Classification on Twitter Data Using Support Vector Machine," Proc. - 2018 IEEE/WIC/ACM Int. Conf. Web Intell. WI 2018, pp. 676–679, 2019, doi: 10.1109/WI.2018.00-13.

[7] J. Žižka, F. Dařena, and A. Svoboda, Text Mining with Machine Learning. 2019.

[8] K. Khan et al., Sentiment Analysis Mining Opinions, Sentiments, and Emotions, vol. 9, no. 9. 2018.

[9] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, "Sentiment analysis of social media Twitter with a case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm," Procedia Comput. Sci., vol. 161, pp. 765–772, 2019, doi: 10.1016/j.procs.2019.11.181.

[10] RapidMiner, "RapidMiner," 2020. https://rapidminer.com/us/ (accessed Nov. 30, 2020).

[11] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, "Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking," IEEE Access, vol. 8, pp. 90847–90861, 2020, doi: 10.1109/ACCESS.2020.2994222.

[12] Z. Zhu, J. Liang, D. Li, H. Yu, and G. Liu, "Hot Topic Detection Based on a Refined TF-IDF Algorithm," IEEE Access, vol. 7, pp. 26996–27007, 2019, doi: 10.1109/ACCESS.2019.2893980.

[13] Help.instagram.com, "Instagram," 2020. https://help.instagram.com/424737657584573 (accessed Oct. 20, 2020).

[14] Tableau, "Tableau," 2021. https://www.tableau.com/products (accessed May 09, 2021).

[15] C. Murkute and S. D. Agashe, "Development of comprehensive decision support system for the binary batch distillation column pilot plant using PYTHON," Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019, vol. 2019-April, no. Icoei, pp. 613–617, 2019, doi: 10.1109/icoei.2019.8862545.