Analysis of Sexual Harassment Tweet Sentiment on Twitter in Indonesia using Naïve Bayes Method through National Institute of Standard and Technology Digital Forensic Acquisition Approach

Kholiq Budiman [1*], Nurul Zaatsiyah [1], Ulfatun Niswah [1], Fikry Muhanna Nur Faizi [1]

[1] Department of Computer Science, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Semarang, Indonesia

*Corresponding author: kholiq.budiman@mail.unnes.ac.id

ARTICLE INFO

ABSTRACT

In this era, the internet is commonly used in society, especially for social media. Social media is an open and unlimited medium of communication where people can freely express their opinions. However, nowadays, many people are abusing social media to do negative things; online sexual harassment is an example. This research was conducted based on previous research or literature studies; the research results showed various sentiment analysis results using the Naïve Bayes classifier method through the digital forensic acquisition approach. The study aims to identify sexual harassment by using sentiment analysis from Twitter and measuring grouping results' accuracy with the Naïve Bayes method. In this study, sexual harassment was identified using analytical sentiment by testing 300 tweets data consisting of 30 queries from Twitter. The result of sentiment analysis from 300 Twitter scraping data shows that the value of negative sentiment is higher than positive sentiment with an average total of 69.7%. From this sentiment, Twitter becomes one of the free communication media and vulnerable to verbal sexual harassment through tweets in Indonesia.

1 Introduction

The development of technology today is progressing very rapidly with technology's influence in various aspects of life. Technological advances are inseparable from the internet's influence; the rapid development of the internet is a factor of ease of access to communication ranging from mobile phones *to smartphones.* The facilities provided *by smartphones* are diverse, one of which is social media. Along with society's development, began to often use social media in everyday life because of the ease of internet access that facilitates communication and does not know remotely or closely.

In this era, the internet is a natural thing to be used by people of this world, especially for those who like to play social media. Social media is a medium of communication for the public to be free in expressing their opinions. Twitter social *media* is one of the means of expressing opinions freely. Social media is growing relatively high, and it can be seen from the increasing number of social media users. Social media is quite important nowadays. According to the hosting rating site, Twitter user data reaches a total of 1.3 billion accounts with 330 million active users and 500 million tweets sent every day. The data proves the public's dependence on social media. Social media is used to greet each other and share stories; now, social media is also used for business. If there is sexual harassment in social media, it will also hinder existing business processes. Kholiq (2020) suggests that businesses must continue to run; whatever happens, there needs to be a backup system to solve the problem.

Twitter as a medium to socialize, discuss, or give each other reviews about an existing phenomenon. Twitter can be used as a medium to express the public opinion that is transparent.

Nevertheless, nowadays, many people are abusing social media to do harmful things. Acts of social media abuse today have occurred a lot in the community. Cyberstalking, cyberbullying, and sexual harassment online are common in the community (Sandes, 2018).

Sexual harassment is an act that often occurs in the community and hurts the victims. A person who has experienced such treatment as sexual abuse by those around him will provide psychological trauma and negatively affect his personality formation. Sexual harassment can be sexual content, making jokes that lead to sexuality and contempt for one's body parts, and making physical contact in the form of touch or the like. Data from the National Commission on Violence against Women (Komnas Perempuan) throughout 2019, there are 239 reports of sexual harassment. The Komnas Perempuan reported that in the period 2011 to 2019, there are 46,698 cases of sexual harassment in public spaces. Data from the Komnas Perempuan there are 23,021 cases of rape, as many as 9,039 cases, sexual harassment, 2,861 cases, and crimes over the internet as many as 91 cases. Data from the Komnas Perempuan recorded a significant increase, namely complaints of cybercrime cases 281 cases (2018 recorded 97 cases) or up by 300%. Most cyber cases are in the form of threats and intimidation of the spread of victim porn photos and videos.

Dealing with such sexual harassment over social media, Indonesia issued Law No. 11 of 2008 on ITE and has been amended by Law No. 19 of 2020 on Amendment to Law No. 11 of 2008 on Information and Electronic Transactions (ITE Law). In LAW ITE Article 27 paragraphs (1) to (4) states that prohibited acts are intentional acts and send and disseminate Electronic Information or Electronic Documents containing violations of decency, gambling, humiliation and/or defamation, extortion and/or intimidation.

Tweet classification testing is done by measuring the accuracy, precision, recall, and f1-score of the Naïve Bayes classifier's calculation results. In research on sentiment analysis of Indonesian tweets with deep belief networks (Winarko, 2017), producing a valid sentiment classification model can use a deep belief network method, Naïve Bayes, and support vector machine. The research obtained results that the calculation of classification with the method Naïve Bayes resulted in a better overall test value than other classification methods. Previous research has discussed the analysis of sentiment regarding the automotive car market on Twitter using Naïve Bayes' method (Rustiana, 2017). Moreover, this study aims to find out the most sold and widely discussed car brands on Twitter. In Rustiana's study, Naïve Bayes' accuracy was obtained at 93%.

This study was conducted to see positive and negative sentiments on all Twitter user tweet test data in Indonesia, leading to sexual harassment. The sentiment analysis results can warn Twitter users to be wise in uploading tweets because the digital footprint will always exist. Especially if in cases of sexual harassment that can be used as evidence of trial if there are parties who feel harassed.

Sexual harassment itself is classified as a cybercrime. Cybercrime can be defined as unlawful acts by utilizing computer technology connected to internet networks (Naufal & Jannah, 2012). In other literature, cybercrime is often likened to computer crime or computer crime. Computer crime itself is a crime against individuals and organizations where the perpetrators use computer technology to commit crimes (Karthikeyan, Geetha, Shiva, & Sivaraj, 2017). Abuse can be done verbally or through physical acts. In terms of online interactions, harassment can occur through inappropriate sentences sent via digital messages. Sending digital messages with content that is immoral can also be categorized as an act of harassment.

Internet harassment knows no bounds because it allows people to interact with each other with or without permission (Zainudin, Hasbullah, Wahab, & Ramli, 2016). Based on reports of online harassment, digital violence, and stalking in America published by the Center for Innovative Public Health Research in 2016, defining harassment can include a wide range of unwanted contacts to create an intimidating environment for victims (Lenhart, Ybarra, Zickuhr, & Price-Feeney, 2016). Online harassment can also be cyberstalking, such as stalkers can monitor and contact their victims through several digital media. According to Danah Boyd, social media's characteristics change such interactions' dynamics; the ability to upload comments anonymously on one's idea can affect online space's culture and context (Octora, 2019). Besides, many forms of online social media abuse can occur at any time, accessible and connected to the victim.

2 The Proposed Method/Algorithm

2.1 Sentiment Analysis

Data obtained from Polda Metro Jaya almost every day shows as many as 25 cases of online harassment on social media. On the other hand, the legal process for handling online harassment cases in Indonesia is still relatively slow. The victims do not report cases of abuse experienced to the police n. Data from national Pew Research in 2017 found that 41% of Americans have been targeted for personal online harassment behavior, and about (66%) have witnessed online harassment behavior towards others. Instagram, Facebook, Twitter is a social media platform widely used for online harassment (Rosyidah & Nurdin, 2018). Often this behavior targets personal or physical characteristics; 14% of Americans say they have been harassed online specifically because of their politics, while one in ten people have been subjected to abuse because of their physical appearance (9%), race, or ethnicity (8%) and gender (8%). According to Saeidi (2019), who classified harassment tweets online, sexual harassment is more comfortable to detect in tweets because of the unequivocal speech was written on the post.

In detecting sexual harassment, a sentence on Twitter must be analyzed. First, the analysis is used to know feelings and opinions based on words that are considered harassing. Sentiment analysis can find opinions from the author about a particular entity (Feldman, 2013). While. According to Tang (Haddi, Liu & Shi, 2013), sentiment analysis on the internet can be used to determine opinions or feelings and responses to a product, incident, or issue. Sentiment analysis is an investigation of public opinion on a topic or event.

Once a sentence is found leading to sexual harassment, it must be classified, whether it is a swear word or just a joke. Naïve Bayes is a machine learning algorithm whose classification efficiency is proven in applications such as document categorization and email spam filtering (Bužić & Dobša, 2018). Naïve Bayes classifier algorithm is an algorithm used to determine the highest probability value that groups test data with the correct category. In this study, the test data is in the form of tweet documents. The classification of documents is divided into two stages. The first stage is training on known category documents. At the same time, the second stage is the classification of documents, whose categories are unknown.

2.2 Digital Forensics Acquisition

Data acquisition in the sense that it is physically or remotely do the retrieval of existing data on the computer. Data is retrieved around all networks connected to a computer or system, with the National Institute of Standard and Technology (NIST) acquisition approach. NIST SP800-30 is a risk management framework used widely to identify, evaluate, and deliver solutions to risks that may arise in using information technology. Sentiment analysis can determine the suspect or perpetrator of tweets that lead to sexual harassment (Riadi, Yudhana, Caesar, & Putra, 2020).

3 Method

This research was conducted based on previous research or literature studies; the research results showed mixed sentiment analysis results using the Naïve Bayes classifier method through the digital forensic acquisition approach. Sentiment analysis is a method of identifying an opinion, assessment, and assessment related to a particular service, event, and topic (Liu, 2012). Sentiment analysis is done by determining an opinion or assessment on an event or service that has a positive or negative value that will later be used as a guideline in improving services (Nurjanah, Perdana, & Fauzi, 2017).

3.1 Population and Sample

In this study, the population and samples used to retrieve data from Twitter using a collection of queries grouped as a word that could become sexual harassment online. This study used 300 tweet data from 30 queries to be processed with an analysis sentiment application for tweet labeling. The study collected data collected by scraping tweets with the Twitter API and acquisition time, as shown in Table 1.

Table 1. Data acquisition time

No	Process	Start Time	End Time
1.	*Scraping Twitter*	27-12-202016:57:29 WIT	28-12-2020 20:47:07 WIB
2.	*Labeling*	27-12-2020 18:49 WIB	28-12-2020 9:30 PM EDT
3.	Accuracy analysis	28-12-2020 10:10 PM EDT	28-12-2020 10:11 PM EDT

3.2 Research Flow and Framework

Sentiment analysis is a method used to analyze the opinions, assessments, evaluations, attitudes, opinions, emotions, and responses of a group of people regarding a particular service, topic, and event. Sentiment analysis is carried out to determine trends related to opinions, comments, positive or negative responses to be used as guidelines in improving the service of rules or public policy. (Nurjanah, Perdana, & Fauzi, 2017). This research uses the previous research library literacy method as the basis for conducting current research and digital forensic acquisition approach in data retrieval (Tarmizi, Saee & Ibrahim, 2020). The digital forensic data acquisition approach is carried out for thorough data retrieval to observe relevant and valid evidence to be used as evidence in a legal case issue.
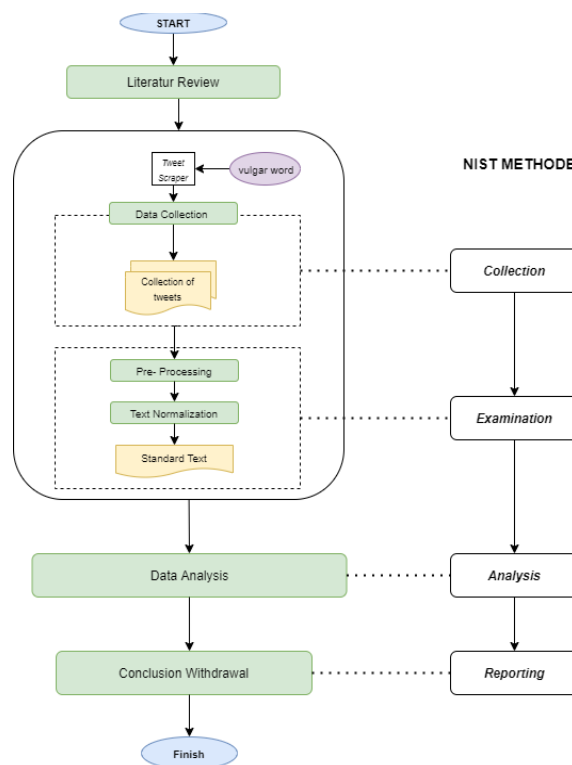
Figure 1. Research framework

3.3 Scrapping Tweet

3.3.1 *Collection*

Stage of collection, identification, label, record, and retrieve data from the source. This data is related to the following data integrity maintenance procedures (Umar, Riadi, & Zamroni, 2018).

In this study, the data was obtained through a data set from Twitter globally using a specific query that is sexual harassment. The question word is used to acquire data from public sentiment about sexual harassment on Twitter in Indonesia.

3.3.2 Examination

Data that has been collected for forensics will be processed during the inspection stage. Combine various scenarios (automatic and manual), evaluation, and Maintain data integrity and release the data as needed.

3.3.2.1 Text preprocessing

In this study, text mining is the initial stage. This stage includes the process of preparing data to be used in a text mining system. The steps taken at this stage are LowerCase (converting all characters and letters into lowercase and tokenize), i.e., decrypting the initial description in the form of a sentence into words and deleting the processed separator, such as period (.), comma (,), space, and numeric characters in the word. The next stage after data collection is data preprocess, which includes the following processes:

3.3.2.2 Case folding

The case folding method facilitates the dataset analysis process and reduces memory usage when collecting data through API.

3.3.2.3 Normalization

Ineffective use of sentences is often used when interacting and communicating with other social media users. Ineffective words are formed because human interaction habits are not following language dictionary standards. In this case, the omission of inappropriate words is very influential when data analysis is performed.

3.3.2.4 Data labeling

This method combines qualitative analysis and empirical validation by utilizing human assessment sentiment. Polarity assessment using the lexical dictionary feature with a sentiment score of 5 additional criteria, namely exclamation point, uppercase, word-setting level, polarization shift due to the word "but" and using the tri-gram feature to check the existence of negation.

3.4 Analysis

The analysis stage uses the same method to analyze the results of the examination. For later obtained information and answers that are technically and legally useful for collection and examination. The data analysis process is done to convert text data obtained on Twitter into information due to data collection. The data set has measured the level of accuracy using the method of Naïve Bayes. To then be concluded the data set. According to Xue, Wei, and Guo (2020), the Naïve Bayes formula is shown in Equation 1.

$$P\left(C_j|W_i\right) = \frac{P(C_j) \times PW_i|C_j)}{P(W_i)} \tag{1}$$

The author tries a sentence

Sentence = "*autisssssss pacar udah nggak perawan sayang terima aja jadi*"

$$P(Positif|Kalimat) =$$

$$\frac{\begin{matrix} P(Positif|autisssssss) \times P(Positif|pacar) \times P(Positif|udah) \times P(Positif|nggak) \times \\ P(Positif|perawan) \times P(Positif|sayang) \times P(Positif|terima) \times P(Positif|terima) \times \\ P(Positif|aja) \times P(Positif|jadi) \times P(Positif) \end{matrix}}{\begin{matrix} P(autisssssss) \times P(pacar) \times P(udah) \times P(nggak) \times P(perawan) \times P(sayang) \\ \times P(terima) \times P(aja) \times P(jadi) \end{matrix}} \tag{2}$$

$$P(Positif|Kalimat) = \frac{1 \times 0.287 \times 0.7 \times 0.75 \times 0.23 \times 0.5 \times 0.67 \times 0.458 \times 0.142 \times 0.366}{1 \times 0.033 \times 0.047 \times 0.018 \times 0.061 \times 0.0142 \times 0.113 \times 0.047} \tag{3}$$

$$P(Positif|Kalimat) = 2136647.75345 \tag{4}$$

$$P(Negatif|Kalimat) =$$

$$\frac{\begin{array}{c} P(Negatif|autissssss) \times P(Negatif|pacar) \times P(Negatif|udah) \times P(Negatif|nggak) \times \\ P(Negatif|perawan) \times P(Negatif|sayang) \times P(Negatif|terima) \times P(Negatif|terima) \times \\ P(Negatif|aja) \times P(Negatif|jadi) \times P(Negatif) \end{array}}{\begin{array}{c} P(autissssss) \times P(pacar) \times P(udah) \times P(nggak) \times P(perawan) \times P(sayang) \\ \times P(terima) \times P(aja) \times P(jadi) \end{array}} \tag{5}$$

$$P(Negatif|Kalimat) = \frac{1 \times 0.857 \times 0.4 \times 0.5 \times 0.846 \times 0.666 \times 0.666 \times 0.583 \times 10.639}{1 \times 0.033 \times 0.047 \times 0.018 \times 0.061 \times 0.0142 \times 0.113 \times 0.047} \tag{6}$$

$$P(Negatif|Kalimat) = 18560745.338 \tag{7}$$

$$P(Positif|Kalimat) < P(Negatif|Kalimat) \tag{8}$$

3.5 Reporting

Report the results of the analysis, a description of the actions taken, explain the tools and procedures selected and determine other actions required in execution (for example, forensic examination of other data sources, identify gaps or strengthen existing security controls), and provide suggestions for improvements, procedures, devices, and other aspects of forensic procedures.

4 Results and Discussion

The acquisition process was done by preparing a data query, which was used to scrape Twitter data. The query that has been entered is useful to see if a tweet can be a case of sexual harassment online or not.
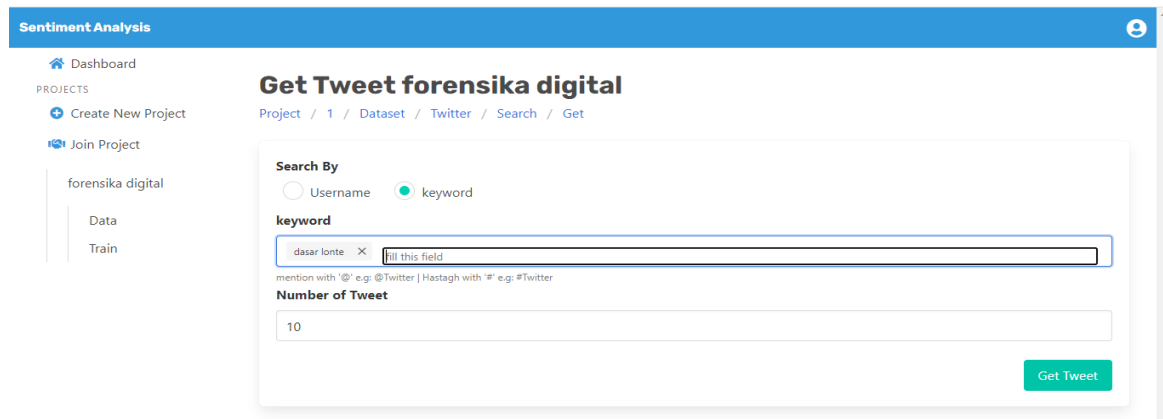

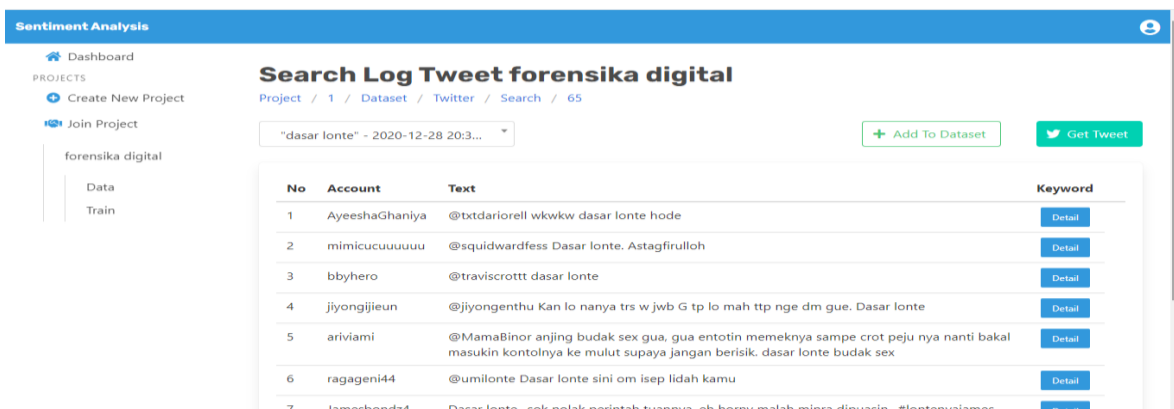**Figure 2.** Data collection scraping process tweets

Figure 3. Data collection results

At this stage, evidence of tweet data obtained from the query that has been entered. Data and physical evidence will be recorded and stored for later data labeling to know the sentence class and Naïve Bayes' needs, as seen in Figure 4.

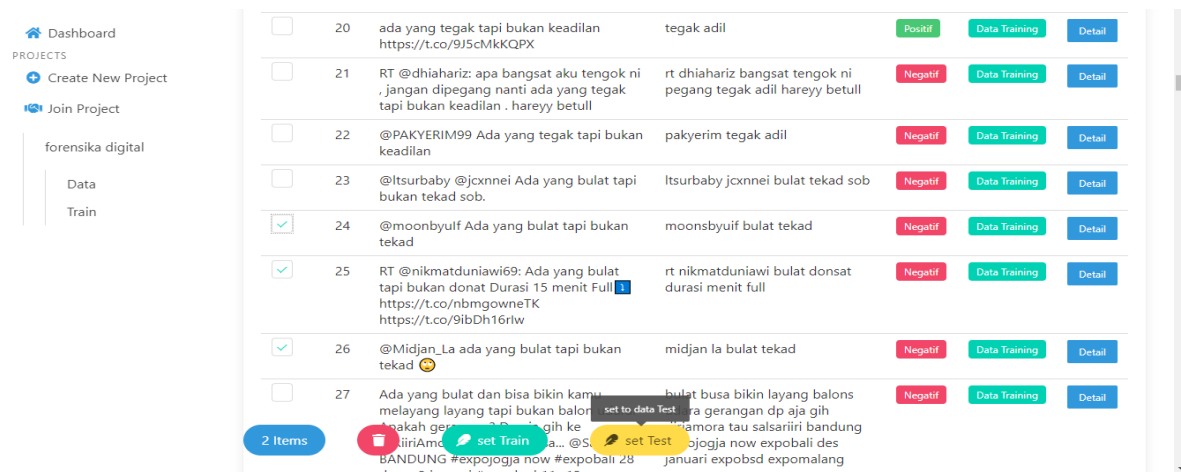**Figure 4.** The process of labeling tweet data (text processing)

Figure 5. Data testing grouping process

The data collection results were divided into two parts with a composition of 70%:30%, where 70% is used for training data and 30% for data testing. Training data were processed as a reference on the method Naïve Bayes. The training data ware validated with data testing to get confusion matrix results from the training process, as shown in Figure 6.
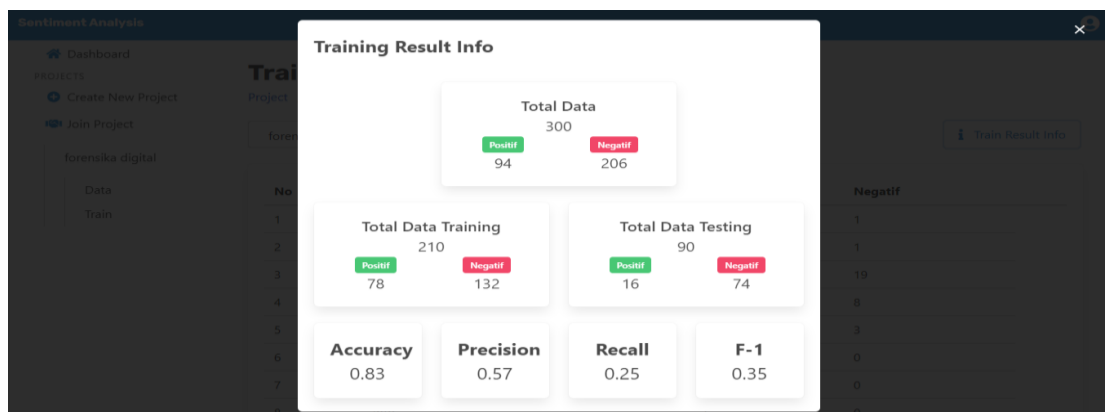

Figure 6. Analysis of sentiment results

The study aims to identify sexual harassment by using sentiment analysis from Twitter and measuring grouping results' accuracy with the Naïve Bayes method. The results of this research can be used as a reminder for Indonesian people to use social media appropriately, wisely, and responsibly. In this study, sexual harassment was identified using analytical sentiment by testing 300 tweet data consisting of 30 queries from Twitter. The results of such tests are shown in Table 2.

Table 2. Sentiment analysis results

Description	Total Data	Positive	Negative
Data *Scraping*	300	94 31,3%	206 69,7%
Training data	210	78	132
Data *Testing*	90	16	74

From the sentiment analysis results, the average value of negative sentiment analysis is higher than the analysis of positive sentiment. This research indicates that the keyword data indicated to lead to sexual harassment on Twitter is used by Twitter users in Indonesia to tweet negative value things.

The results of sentiment analysis with the Naïve Bayes method are designed to determine the data sets' accuracy. Test results using this method will result in accuracy, precision, recall, and F1 values. Table 3 shows the results of the Naïve Bayes classification test.

Table 3. Naïve Bayes test results

Accuracy	*Precision*	*Recall*	*F-1*
83%	57%	25%	35%

Based on Table 3, it is known that the accuracy of classification calculation with the Naïve Bayes method is 83%, precision 57%, recall 25%, and F1-score 35%. It can be concluded from the results of sentiment analysis and testing method Naïve Bayes, and it is known that the use of Twitter freely and inappropriately makes Twitter users in Indonesia dare to upload tweets that lead to acts of harassment, primarily verbal sexual abuse. The analysis results and direct observation of tweet data found that Twitter becomes a means of possible verbal sexual abuse such as vulgar communication. A solicitation to have intercourse, sexual comments about a person's body, sexually dirty jokes to demean others' physical. According to the data, positive sentiment referred to positive tweets containing warnings and criticisms to tweets indicated as sexual harassment.

The dissemination of pornographic-related information, photos, and videos, including sexual harassment, can be found easily on Twitter. The use of Twitter in Indonesia can freely make Twitter social media a medium that can increase access to pornography to sexual harassment. Such an increase can trigger the occurrence of sexual crimes and various other irregularities. This study's results remind Twitter users in Indonesia to be wiser and more careful, and critical in using one of the world's largest communication media. Because with negative tweets that lead to sexual harassment, both offensive to individuals and groups, whatever the reason, the tweet's uploader must be held accountable if the victim takes legal action for the abuse.

5 Conclusion

Based on the research results, it can be concluded that the study used 30 queries related to sexual harassment sentences to obtain data from Twitter. Sentiment analysis results from 300 Twitter memo data showed that the value of a negative sentiment is higher than the positive view, with a total average

of 69.7%. From this sentiment, it is understood that Twitter social media is a free communication medium and vulnerable to verbal sexual abuse in Indonesia through tweets. Twitter is an easy-to-use tool that accesses all kinds of information from positive to negative. The lack of strict filters and policies to spread negative information (mostly related to sexual harassment) makes Indonesian Twitter users unwise in using negative information. This research uses the Naïve Bayes method to conduct classification and determine the accuracy with test results of 83% overall accuracy value.

References

Budiman, K., Arini, F. Y., & Sugiharti, E. (2020). Disaster Recovery Planning with Distributed Replicated Block Device in Synchronized API Systems. In *Journal of Physics: Conference Series,* 1567(3), 032023. IOP Publishing. doi:10.1088/1742-6596/1567/3/032023

Bužić, D., & Dobša, J. (2018). Lyrics classification using Naïve Bayes. In *2018 41st International Convention on Information and Communication Technology, Electronics, and Microelectronics (MIPRO)*, 1011-1015. IEEE. doi:10.23919/MIPRO.2018.8400185

Feldman, R. (2013). Techniques and Applications for Sentiment Analysis. *Communications of the ACM*, *56*(4), 82-89. doi:10.1145/2436256.2436274

Haddi, E., Liu, X., & Shi, Y. (2013). The Role of Text Pre-processing in Sentiment Analysis. *Procedia Computer Science*, *17*, 26-32. doi:10.1016/j.procs.2013.05.005

Karthikeyan, R., Geetha, D. T., Shiva Prakash, S., & Sivaraj, S. (2017). Strategy of Triple-E on Solving Trojan Defense in Cyber Crime Cases. *International Journal for Research & Development in Technology*, *7*(1), 167-171. Retrieved from http://bowringj.people.cofc.edu/classes/csci%20392/docs/sdarticle.pdf

Lenhart, A., Ybarra, M., Zickuhr, K., & Price-Feeney, M. (2016). Online harassment, Digital Abuse, and Cyberstalking in America. *Data and Society Research Institute*. Retrieved from https://datasociety.net/library/online-harassment-digital-abuse-cyberstalking/

Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis lectures on Human Language Technologies*, *5*(1), 1-167. doi:10.2200/S00416ED1V01Y201204HLT016

Naufal, M. M., & Jannah, H. S. (2012). Penegakan Hukum Cyber Crime ditinjau dari Hukum Positif dan Hukum Islam [Cyber Crime Law Enforcement in terms of Positive Law and Islamic Law]. *Al-Mawarid Journal of Islamic Law*, *12*(1), 42565. Retrieved from https://www.neliti.com/publications/42565/penegakan-hukum-cyber-crime-ditinjau-dari-hukum-positif-dan-hukum-islam

Nurjanah, W. E., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen terhadap Tayangan Televisi berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet [Sentiment Analysis on Television Shows based on Public Opinion on Twitter Social Media using the K-Nearest Neighbor Method and Weighting the Number of Retweets]. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer,* 1(12), 1750-1757. Retrieved from http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/631

Octora, R. (2019). Problematika Pengaturan Cyberstalking (Penguntitan di Dunia Maya) dengan menggunakan Anonymous Account pada Sosial Media [The Problems of Cyberstalking Settings (Cyberspace Stalking) Using Anonymous Accounts on Social Media]. *Dialogia Iuridica: Jurnal Hukum Bisnis Dan Investasi*, *11*(1), 77-96. doi:10.28932/di.v11i1.1902

Riadi, I., Yudhana, A., Caesar, M., & Putra, F. (2017). Analisis Recovery Bukti Digital Instagram Messengers menggunakan Metode National Institute of Standards and Technology (NIST) [Instagram Messengers Digital Evidence Recovery Analysis using the National Institute of Standards and Technology (NIST) Method]. In *Prosiding Seminar Nasional Teknologi Informasi dan Komunikasi (SEMANTIKOM)*, 161-166. Retrieved from https://semantikom.unira.ac.id/2017/SEMANTIKOM_2017_paper_26.pdf

Rosyidah, F. N., & Nurdin, M. F. (2018). Media Sosial: Ruang Baru dalam Tindak Pelecehan Seksual Remaja [Social Media: A New Space for Adolescent Sexual Harassment]. *Sosioglobal: Jurnal Pemikiran dan Penelitian Sosiologi*, *2*(2), 38-48. Retrieved from http://jurnal.unpad.ac.id/sosioglobal/article/view/17200

Rustiana, D., & Rahayu, N. (2017). Analisis Sentimen Pasar Otomotif Mobil: Tweet Twitter Menggunakan Naïve Bayes [Automobile Automotive Market Sentiment Analysis: Tweet Twitter Using Naïve Bayes]. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, *8*(1), 113-120. doi:10.24176/simet.v8i1.841

Saeidi, M., Sousa, S. B. D. S., Milios, E., Zeh, N., & Berton, L. (2019, September). Categorizing Online Harassment on Twitter. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 283-297. Springer, Cham. doi:10.1007/978-3-030-43887-6_22

Sandes, T. M. (2018). *Klasifikasi Tweet Pelecehan Online pada Twitter dengan menggunakan Metode Naïve Bayes [Online Classification of Harassment Tweets on Twitter using the Naïve Bayes Method]* (Doctoral dissertation). Retrieved from http://repository.uin-suska.ac.id/16882

Tarmizi, N., Saee, S., & Ibrahim, D. H. A. (2020). Detecting the Usage of Vulgar Words in Cyberbully Activities from Twitter. *International Journal on Advanced Science, Engineering and Information Technology*, 10(3), 1117-1122. doi:10.18517/ijaseit.10.3.10645

Umar, R., Riadi, I., & Zamroni, G. M. (2018). Mobile Forensic Tools Evaluation for Digital Crime Investigation. *International Journal on Advanced Science, Engineering and Information Technology*, *8*(3), 949. doi:10.18517/ijaseit.8.3.3591

Winarko, E. (2017). Sentimen Analisis Tweet Berbahasa Indonesia dengan Deep Belief Network [Sentiment Analysis of Indonesian Language Tweets with the Deep Belief Network]. *Indonesian Journal of Computing and Cybernetics Systems (IJCCS)*, *11*(2), 187-198. doi:10.22146/ijccs.24716

Xue, Z., Wei, J., & Guo, W. (2020). A Real-Time Naïve Bayes Classifier Accelerator on FPGA. *IEEE Access*, *8*, 40755-40766. doi:10.1109/ACCESS.2020.2976879

Zainudin, N. M., Zainal, K. H., Hasbullah, N. A., Wahab, N. A., & Ramli, S. (2016, May). A Review on Cyberbullying in Malaysia from Digital Forensic Perspective. In 2016 *International Conference on Information and Communication Technology* (ICICTM), 246-250. IEEE. doi:10.1109/ICICTM.2016.7890808