

Sentiment Analysis on the Ratification of Penghapusan Kekerasan Seksual Bill on Twitter

<http://dx.doi.org/10.28932/jutisi.v7i3.4051>

Riwayat Artikel

Received: 07 Oktober 2021 | Final Revision: 25 November 2021 | Accepted: 26 November 2021

Gary Dimitri Hamidi[✉]#1, Farida Afira Bestari^{#2}, Alexandra Situmorang^{#3},
Nur Aini Rakhmawati^{#4}

Information Systems Department, Institut Teknologi Sepuluh Nopember
Kampus ITS Sukolilo, Surabaya - 60111, Jawa Timur, Indonesia

¹dimitri.18052@mhs.its.ac.id

²afira.18052@mhs.its.ac.id

³situmorangalexandra.18052@mhs.its.ac.id

⁴nur.aini@is.its.ac.id

Abstract — *Rancangan Undang-Undang Penghapusan Kekerasan Seksual (RUU PKS)*, or the bill of the Republic of Indonesia on the Elimination of Sexual Violence, is a bill that discusses sexual violence, victim protection and its scope concern the matters related to sexual violence. Elimination of sexual violence according to the PKS Bill aims to prevent all forms of violence. These discussions and conversations also occur on social media, especially on Twitter. Taking public sentiment is significant in choosing the proper messages, interference, and policy. Sentiment analysis is a field of study that analyses opinions, sentiments, judgments, evaluations of a person attitudes and emotions regarding a particular topic, service, product, individual, organization or activity. This study used the method of crawling to get data from Twitter. Then data cleansing, data processing is carried out using Bernoulli, Logistic Regression, and Support Vector Classification (SVC) algorithm. The data is then evaluated using three methods: accuracy, classification report, and confusion matrix. Based on the three algorithms used, it is found that all methods are equally accurate with 0.65. This study found positive, negative, and neutral sentiments expressed to the bill of Elimination of Sexual Violence through comments. It is shown that most people using the keyword “RUU PKS” are positive to the bill of Elimination of Sexual Violence (RUU PKS) while most people’s sentiments using #RUUPKSBukanSolusi are negative to the bill.

Keywords— classification; sentiment analysis; ruu pks.

I. INTRODUCTION

RUU PKS, or the bill of Republic of Indonesia on the Elimination of Sexual Violence, is a bill that discusses sexual violence, victim protection and its scope concern the matters related to sexual violence. The principles and objectives of the PKS Bill itself are based on respect for human dignity, non-discrimination, the best interests of victims, justice, benefit for all, and legal certainty. Elimination of sexual violence according to the PKS Bill aims to prevent all forms of violence. The PKS Bill also attempts to handle, protect, and recognize victims, take action against perpetrators and create an environment free of sexual violence. The elimination of Sexual Violence includes prevention, treatment, protection, recovery of victims and prosecution of perpetrators. The elimination of sexual violence as referred to in article 4 section 1 is an obligation of the state [1].

According to Deputy Chairperson of the Legislative Body of the Indonesian House of Representatives (DPR) Willy Aditya, the Bill on the Elimination of Sexual Violence has an urgency to be ratified considering the statistical report made by National Commission on Human Rights (KOMNAS HAM) that the amount of violence against women continues to rise significantly [2]. Recently, public attention to sexual violence in Indonesia has also increased. The community began to have vigilance and realized the importance of discussing sexual violence and legal protection for victims. This is increasing public pressure for the ratification of the PKS Bill. The pressure and criticism began when the Indonesian House of Representatives took out

the PKS Bill from the 2020 national legislation program [3]. The national legislation program is a list of priority bills for ratification.

These discussions and conversations did not escape being carried out on social media, especially on Twitter. Twitter users can create a post that contains a limited-character status. These posts are called Tweets. Indonesia has a recorded twitter users of 14.05 million users [4]. Twitter itself can be a medium to convey complaints or ideas, and in this case, particular opinions and urges regarding the bill on the Elimination of Sexual Violence through the keyword of “RUU PKS” and #RUUPKSBukanSolusi.

Sentiment analysis is a field of study that analyzes opinions, sentiments, judgments, evaluations of a person's attitudes and emotions regarding a particular topic, service, product, individual, organization or activity. Sentimental analysis has been used for analyzing public's sentiment. Taking public sentiment is significant in choosing the proper messages, interference, and policy. This technique had been used in research and also for commercial purposes [5]. The usage of Sentimental analysis on social media is a new area of study. Recently the usage of sentimental analysis on social medias are used for various purposes. Social media like twitter is better suited for sentimental analysis than documents that have longer texts [6], [7]. Previous study had used sentimental analysis to analyze public's sentiment regarding to widely discussed matters. A study conducted by Melton et al. did a lexical based sentimental analysis to find out public sentiment related to COVID-19 vaccine. It also divided the topics into several sections and ten words from each combined dataset [8]. A study conducted with the goal of finding public's sentiment on political issues had also been done to analyze and understand public's opinion [9]. A study aiming to conduct a sentimental analysis of RUU PKS using text mining had been done by Khalisa Virra et al. with naïve bayes, decision tree and random forest algorithm. This study however is more focused on the accuracy between three different algorithms and its performance rather than the evaluation of the sentiment [10].

Public opinion and sentiment regarding the ratification of the PKS Bill can be found out through the sentiment analysis method. By knowing public sentiment regarding the PKS Bill, insight into opinions among the community can also be explored. Input on public opinion, especially on Twitter social media about the PKS Bill can also be taken into consideration to find out public opinions and reactions regarding the PKS Bill. In this study, a sentiment analysis will be carried out on the ratification of the PKS Bill through the keyword of “RUU PKS” and hashtag #RUUPKSBukanSolusi.

Only data obtained from Twitter in the form of opinions, information, or recommendations regarding the PKS Bill accompanied by the keyword “RUU PKS” and hashtag #RUUPKSBukanSolusi will be used in this study. The data on the tweets of people's opinions and reactions on Twitter will be processed. The approaches that will be used in this study is the method of logistic regression, Bernoulli, and Support Vector Machine.

The purpose of this study will focus on finding out the public sentiment regarding the PKS Bill and looking for differences in the results of the analysis with accounts that use bots and those that do not. In addition, this study also aims to obtain the best classification algorithm model that can be used to perform sentiment analysis.

The limitation of this research problem is the manual labeling of data which will refer to a dictionary of words with positive, negative, and neutral sentiments. Through the logistic regression, Bernoulli, and Support Vector Machine approaches, the public sentiment towards the ratification of the PKS Bill, the distribution of words used by the public when expressing opinions about the PKS Bill, and the results of the analysis of the difference between using bots and not using bots can be known. The results of this study can be used as a material for consideration of public opinion regarding the ratification of the PKS Bill which is expected to bring justice to the community and survivors as well as protect the victims.

II. RESEARCH METHOD

A. Literature Review

Sentiment analysis is the process of analyzing online writing to determine whether the emotional tone it conveys is positive, negative, or neutral. Sentiment analysis is often performed on textual data to help businesses monitor brand and product sentiment in customer feedback and understand customer needs [11]. Sentiment analysis provides answers into what the most important issues are. Because sentiment analysis can be automated, decisions can be made based on a significant amount of data rather than plain intuition that isn't always right.

Logistic regression is a process of modelling the probability of a discrete outcome given an input variable. The most common logistic regression models a binary outcome; something that can take two values such as true/false, yes/no, and so on. Multinomial logistic regression can model scenarios where there are more than two possible discrete outcomes. Logistic regression is a useful analysis method for classification problems, where you are trying to determine if a new sample fits best into a category. As aspects of cyber security are classification problems, such as attack detection, logistic regression is a useful analytic technique [12].

The reason for using logistic regression in this sentiment analysis is because when applying machine learning methods to the case at hand, we often immediately choose a machine learning method that is popular and has good performance, for example, neural network methods, especially deep learning. In fact, we often have very little data, while deep learning

methods are data hungry methods, meaning that they require a lot of data in training. We should have experimented with simple machine learning methods first and then on to more advanced methods, because it could be that the performance results with simple methods are better. In addition, methods such as deep learning often also require expensive computational tools.

In this study, to classify positive sentiment and negative sentiment we use the Support Vector Machine (SVM). The Support Vector Machine was first introduced by Vapnik in 1992 as a harmonious series of superior concepts in the field of pattern recognition. Support Vector Machine is one of the methods in supervised learning that is usually used for classification (such as Support Vector Classification) and regression (Support Vector Regression). In classification modelling, SVM has a more mature and clearer mathematical concept than other classification techniques. SVM can also solve classification and regression problems with linear and non-linear.

BernoulliNB which stands for Bernoulli Naive Bayes is a family of Naive Bayes classification methods designed for binary/Boolean features. BernoulliNB implements Naive Bayes training and classification algorithms for data distributed according to a multivariate Bernoulli distribution, i.e. there may be several features but each is assumed to be a binary (Bernoulli, Boolean) valued variable. Therefore, this class requires samples to be represented as binary-valued feature vectors; if given another data type, the BernoulliNB instance can binarize its input (depending on the binarization parameters) [13].

In analysing sentiment, there are several steps that must be taken to get optimal results. In this research, the following steps were carried out: Data Pre-Processing, Data Processing, Evaluation, and Implementation.

B. Data Pre-Processing

The stage begins with the collection process by determining keywords that match the problem then the data is extracted from Twitter. The keywords used are: “RUU PKS”, and #RUUPKSBukanSolusi. The data taken is a collection of tweets with related keywords. The data is then saved in the form of .csv. The flow can be seen in Figure 1. Although this figure was made using RapidMiner program, this experiment doesn’t use RapidMiner as a tool in its methodology.

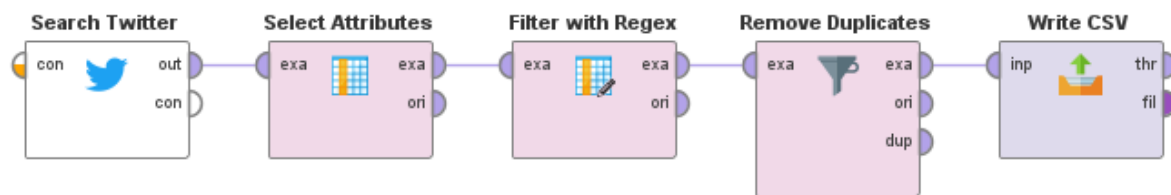


Figure 1. Twitter data crawling scheme.

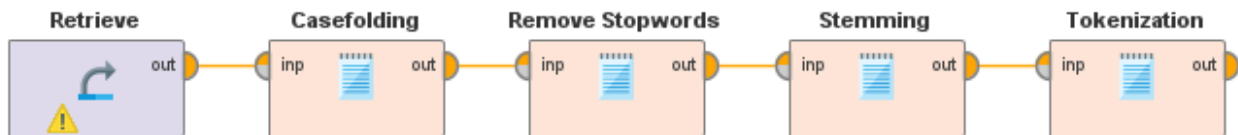


Figure 2. Data cleansing scheme.

The next stage is data cleansing. At this stage, the data is cleaned so that it is ready for analysis. As seen in Figure 2, the steps taken are as follows [14]:

- 1) *Document Filtering*: a process that aims to remove noise in the data such as links, hashtags, and usernames.
- 2) *Case Folding*: the process of changing the text in the document into lowercase.
- 3) *Stop words and Punctuation Removal*: the process of removing words and punctuations that are often used but have no important meaning, such as “is”, “which”, “at” and so on. This stage aims to focus more on important words.
- 4) *Stemming*: the process of breaking down the form of a word into its basic form.
- 5) *Slang Removal*: Removal of slang word is done using a language dictionary. The slang word in each tweet is changed to the standard language found in the slang language dictionary .
- 6) *Tokenization*: the process of separating text into smaller parts into tokens such as words, phrases, or other elements. In this study, tokenization is done by separating sentences into words. The tokenization used in this study is TF-IDF. Unigram is used for the analysis of the distribution of words contained in the document.

C. Data Processing

In this study the data that has been collected is labelled manually by one person. Data labelling is carried out by categorizing the data into 3 classes of sentiment, which are positive, neutral, and negative. Then, bot detection is carried out on each profile that tweets to see if there is a prevalence of certain sentiments if the twitter account is found as a bot later.

Then the performance test of the model is carried out. The models used are models for the supervised learning category because there is already a manual label given to the data. The models used are SVM, Logistic Regression, and BernoulliNB. TF-IDF weighting is used to perform weighting based on the frequency of occurrence of a term in the assembled document and to measure how widely the term is spread in the related document.

D. Evaluation

At this stage, an evaluation of the performance of each model has been carried out. The evaluation used are

- *confusion matrix*, represent the predictions and actual conditions of the data generated by the algorithm,
- *accuracy rate*, represents how many correct predictions, and
- *classification report*, measuring the quality of predictions with a classification algorithm

All three evaluation methods are available in python using the *scikit.learn* library. The evaluation results obtained include accuracy, precision, recall, and F1 values. After obtaining several evaluation results, an analysis of the evaluation results is carried out and determines the best model performance to be used as a tool to classify public sentiment related to the PKS Bill.

E. Implementation

At this stage, the best model that has been obtained after the evaluation stage is carried out on the latest twitter data that discusses the PKS Bill. The data was taken using Apify Twitter Crawler. After the data is obtained and the model is run, the results of the sentiment analysis on the data are obtained. Then, conclusions are made based on the results of sentiment classification using the machine learning classification model.

TABLE 1
DATA PRE-PROCESSING

Step	Before	After
Cleaning Punctuations	Di ujung bulan september, dua tahun lalu, terjadi demo besar di berbagai titik di indonesia. Demo itu menuntut 8 hal. Masih pada ingat nggak?	Di ujung bulan September dua tahun lalu terjadi demo besar di berbagai titik di Indonesia Demo itu menuntut 8 hal Masih pada ingat nggak?
Cleaning URL	Orangtua korban curiga dengan perubahan sikap anaknya yang masih berusia 15 tahun https://t.co/y8fsHlzIFY	Orangtua korban curiga dengan perubahan sikap anaknya yang masih berusia 15 tahun.
Cleaning Emoji / Emoticon	Indonesia mah, ruu pks aja tak kunjung disah'in 😊	Indonesia mah, ruu pks aja tak kunjung disah'in
Cleaning Numbers	Pelaku mengaku telah mencabuli 34 santriwatinya dalam kurun waktu 3 tahun.	Pelaku mengaku telah mencabuli santriwatinya dalam kurun waktu tahun.
Case Folding	Gedeg bgt aku tiap liat berita kekerasan terus pelakunya malah kek gapunya dosa. Itu tau ga sih kalau penyintas bisa trauma sampe sumur hidupnya??	gedeg bgt aku tiap liat berita kekerasan terus pelakunya malah kek gapunya dosa. itu tau ga sih kalau penyintas bisa trauma sampe sumur hidupnya??
Removing Stop Words	pelecehan dan kekerasan seksual yang terjadi di tempat kerja seringkali tidak teridentifikasi dan tidak dilaporkan karena korban mengalami ancaman dari pelaku.	pelecehan kekerasan seksual kerja seringkali teridentifikasi dilaporkan korban mengalami ancaman pelaku.
Stemming	hilangkan stigma terhadap korban dan dorong pengesahan ruu penghapusan kekerasan seksual untuk pencegahan dan penanganan kekerasan seksual bagi semua.	hilang stigma terhadap korban dorong kesah ruu hapus keras seksual untuk cegah tangan keras seksual bagi semua

III. RESULTS AND DISCUSSIONS

The result of the experiments and analysis would be explained in this part. Some results gained would answer the problems defined in previous section. The pre-processing example could be seen in Table 1. This pre-processing dataset would be processed until it could be used as a machine learning classification model.

After the model was trained using the finalised dataset, it would be implemented on the data gathered from Twitter. The amount of training data used from twitter is 460 data. After completion, the following conclusions can be drawn:

1. *The best model for sentiment classification*

The algorithm used for the data are Bernoulli, Logistic Regression and SVM. One predictor variable was vectorized using TF-IDF along with one dependent variable. The weight of TF-IDF is used to determine the relevance of the data to the topic. This weight is used to evaluate how relevant the words mentioned in a corpus. Machine evaluation process was done by paying attention to macro F1, recall and precision variables. From the data mentioned, the evaluation result was obtained as follows in Table 2, with the highest scores shown in bold.

TABLE 2
EVALUATION OF THE PERFORMANCE OF THE MODEL

Machine Learning Model	Precision	Recall	F1
Bernoulli	0.65	0.65	0.65
Logistic Regression	0.65	0.62	0.65
SVM	0.63	0.57	0.58

Evaluation results showed that *Bernoulli* is the best model to for analysing sentiments as it has the highest score of *precision, recall, and F1*.

2. *Sentiment Analysis Result on RUU PKS*

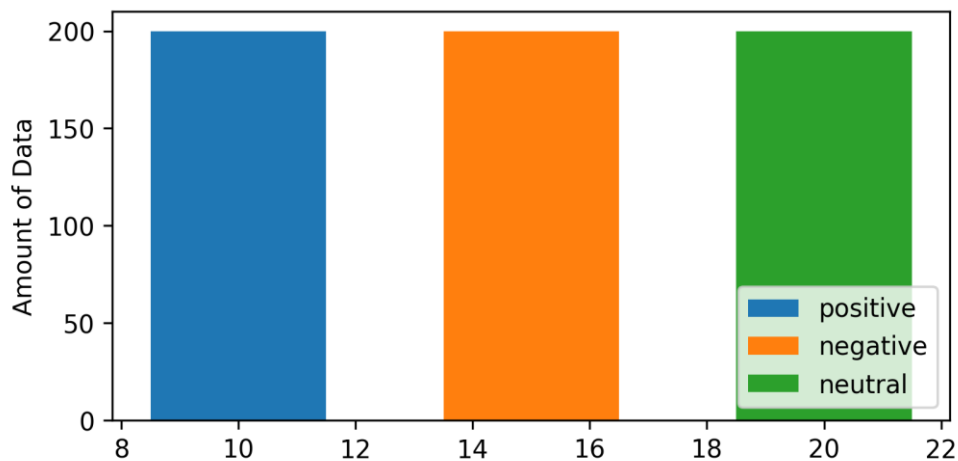


Figure 3. Sentiment proportion of the dataset

The sentiment result is shown in **Figure 3** The blue bar indicates a sentence that is positive, while the orange bar indicates a negative trait, and the green bar indicates a neutral trait. From this figure, it can be seen that the proportion for positive, negative and neutral are 200.

Based on Figure 5, it can be seen from the analysis of positive sentimental tweets that the top 3 words from the word cloud displayed are the words **seksual**, **keras** and **korban**. From the word cloud of positive sentiment tweets, most Indonesians use the word *keras*, *seksual* and *korban* which are referring to *kekerasan seksual* and *korban*. *Kekerasan seksual* translates to sexual violence, and *korban* translates to victims. It can be said that the tweets mainly discuss the victims of sexual violence, which is greatly connected to the bill of sexual violence erasure. Most positive sentiments tweets also drive the urgency of the ratification of RUU PKS, amplifying the effects of not having a law to address sexual crimes by painting the conditions of victims in the tweets. Positive sentiments also express frustrations over the delay of the ratification and most of them rant about the delay and condemn DPR for having to have done so.

The words used in the data for positive sentiment can be seen in **Figure 6**.

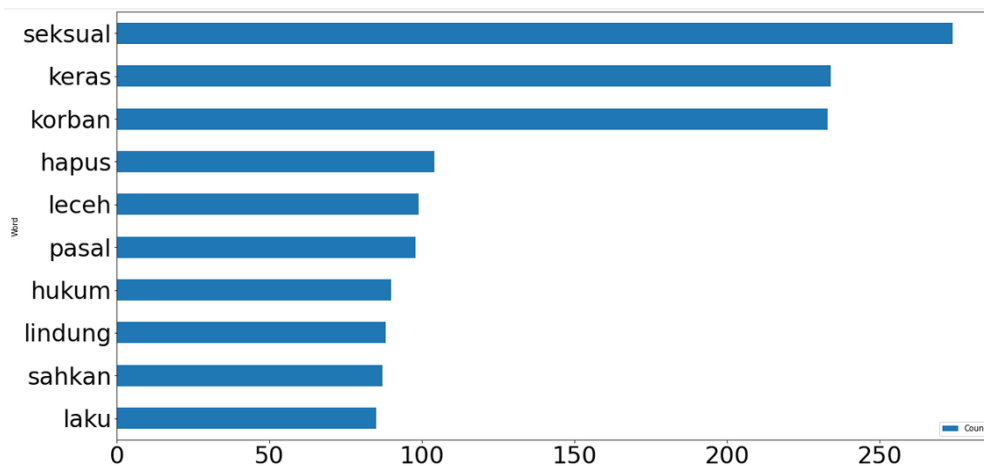


Figure 6. The word distribution of data with neutral sentiment

Figure 6 shows the results of word distribution on positive sentiment data in the form of unigrams (one word). The word *seksual* occupies the top position with a frequency of more than 250 times. The word '*keras*' and '*korban*' comes next with the frequency of more than 200 mentions.

TABLE 3
USE OF THE WORD SEKSUAL

Tweet	Selected Word	Sentiment
ruu pks desak sah ruu atur lindung korban keras seksual	<i>Seksual</i>	Positive
butuh lindung ruu pks penghapusan keras seksual bahas sah	<i>Seksual</i>	Positive
korban kerap sulit lapor keras seksual alami ruu pks sah	<i>Keras</i>	Positive
sahkan ruu pks korban keras seksual harus dapat lindung peran dpr	<i>Keras</i>	Positive
hmmm pasal ruu pks hapus tuh gmn yah kasih hukum laku hak korban dipertimbangkan kayak	<i>Korban</i>	Positive
keras seksual perempuan laki laki juga untuk yuk samasama dukung korban	<i>Korban</i>	Positive

From **Table 3**, it can be perceived that people used the word *seksual* to illustrate positive sentiment. This shows the tendency of the community to support the policy. The word **keras** is used to recognise the bad treatment people show about RUU PKS. The word **keras** is used above 200 times throughout the data. People also utilised the word **korban** for more than 200 times to show their positive sentiment about RUU PKS. This can be taken that RUU PKS has positive attitudes among people who used twitter.



Figure 7. Word cloud for tweets with neutral sentiments

Figure 7 describes the world cloud for neutral sentiment tweets. The top tweets displayed here are *seksual*, *keras*, and *tpks*. The words *seksual* and *keras* are most likely derived from *kekerasan seksual* or sexual violence. TPKS itself is an abbreviation for *Tindak Pidana Kekerasan Seksual* or Sexual Violence Crime, the newest epithet for RUU PKS. It can be inferred from the neutral sentiment tweets that most tweets discuss about the contents of RUU PKS and the plans to change its name to TPKS and the announcement of the change of name. This can be seen from the mentions of *baleg*, *judul*, *ganti*, *nama*, *ubah*, which mean *badan legislasi* or legislation committee, title, change, name, change respectively. This points to the discussion of the change of name of RUU PKS and the discussion of the bill's contents.

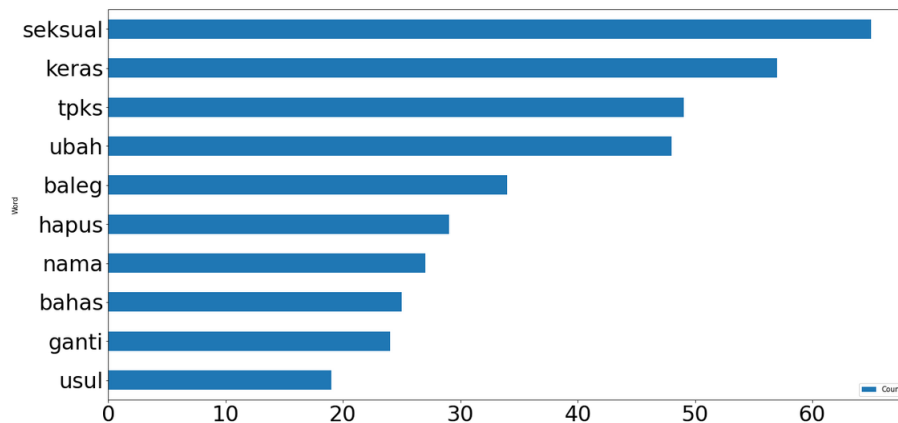


Figure 8. The word distribution of data with neutral sentiments

In Figure 8, the words distribution for neutral sentiment words can be seen. The first position goes to the word *Seksual* or sexual, which gains more than 60 mentions. The second place goes to *keras*, which most likely is derived from the word *kekerasan*, which means violence, with more than 50 mentions. The third position goes to *TPKS*, which is an abbreviation for *Tindak Pidana Kekerasan Seksual*, the new name for RUU PKS.

in life. The tweets have perspective that RUU PKS enable Indonesians to have freedom over sexual actions and therefore repulsing.

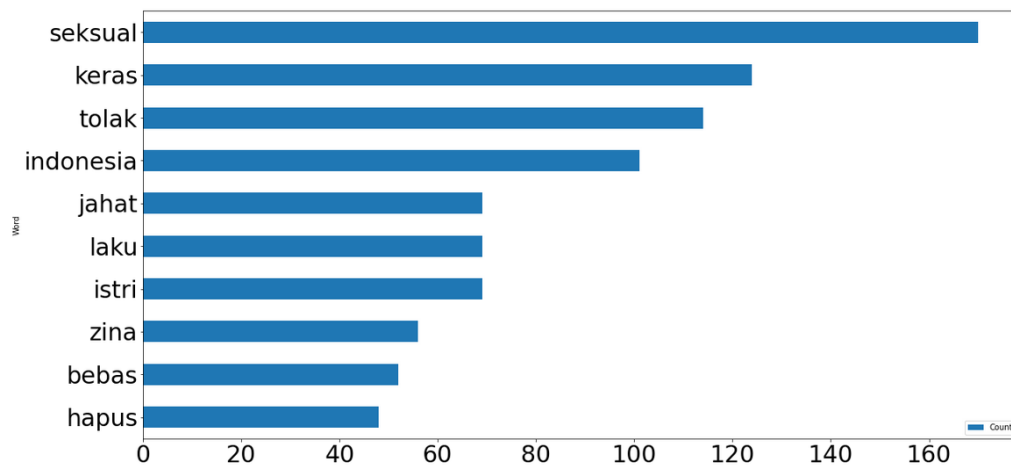


Figure 10. The word distribution of data with negative sentiments

Figure 10 shows the word distribution for negative sentiment words. The word *seksual* which means *sexual* reached over 160 mentions in negative sentiment tweets. The word *keras* comes next, with the mentions over 120 times. The words *seksual* and *keras* are most likely derived from *kekerasan seksual* which means sexual violence. The word *tolak* was placed third, with more over than 100 mentions. *Tolak* means to reject. In this case, the rejection refers to rejection over the ramification of the bill of sexual violence erasure (RUU PKS).

TABLE 5
USE OF THE WORD SEKSUAL

Tweet	Selected Word	Sentiment
Speak up rem kamu bayang orang korban keras seksual kuat bukti dunia	<i>Seksual</i>	Negative
Indonesia butuh uu lindung keras seksual mari dukung kawal bersama dukung ruu pks	<i>Seksual</i>	Negative
ruu pksbukanlah solusi keras pada perempuan anak	<i>Keras</i>	Negative
tolak rancang undangundang hapus keras seksual ruu pks	<i>Keras</i>	Negative
persis ruu pks wajib tolak potensi langgar norma agama	<i>Tolak</i>	Negative
tolak ruu profesional zina bantu menandatangani petisi	<i>Tolak</i>	Negative

From **Table 5**, it can be perceived that people used the word *seksual* to illustrate negative sentiment. The word **keras** is used to recognise the negative treatment people show about RUU PKS. The word **keras** is used above 120 times throughout the data. People also utilised the word **tolak** for more than 110 times but under 120 times to show their negative sentiment about RUU PKS. This can be taken that RUU PKS has positive attitudes among people who used twitter.

IV. CONCLUSION

In this study, a sentimental analysis regarding the ramification RUU PKS on Twitter is done. The analysis is carried out with SVM, Logistic regression and Bernoulli algorithm. The data that is used is taken from 27 January 2019 to 2 October 2021. This study had discussed the positive, neutral, and negative sentiment of RUU PKS. This study has also built a word cloud on the words utilised in people's tweets.

From the examination, it can be said that the most word used to express positive sentiment on RUU PKS are *seksual* with a frequency of more than 250 times, the word *keras* which is mentioned between 200 and 250 times, and *korban* which is also used between 200 and 250 times.

The best model that can be used is Bernoulli with 65% accuracy rate, 0.65 precision score and 0.65 recall score and 0.65 as the F1-Score. Predictor variable used is the sentence on a tweet that has been vectorized by TF-IDF. Whereas the dependent variable is the sentiment determined based on the sentence on the tweet. The weighting scheme variation is done based on queries with macro-F1, recall and precision variables. The selection of the best model is done based on accuracy, precision, and recall. The accuracy level between information is 0.65.

It can be concluded that most Indonesian people are passionate about the ratification of RUU PKS. This result is derived from judging the words and sentences they used to express their opinions regarding RUU PKS, with most agreeing and waiting for the ratification of the bill. Indonesian's tweets also conveyed anger on the justice of sexual crimes and offered sympathy and solidarity for the victims of sexual crimes, opting for the ratification of RUU PKS to give justice for Indonesian society.

The result of this sentiment analysis showed the tendency of Indonesian people in giving positive attitude about the ratification of RUU PKS.

REFERENCES

- [1] DPR, "Rancangan Undang-Undang Republik Indonesia Nomor... Tahun ... Tentang Penghapusan Kekerasan Seksual," 1 2 2017. [Online]. Available: <https://www.dpr.go.id/doksileg/proses2/RJ2-20170201-043128-3029.pdf>. [Accessed 25 10 2021].
- [2] Badan Legislasi, "Angka Kekerasan Meningkat, RUU PKS Harus Segera Disahkan," 16 3 2021. [Online]. Available: <https://www.dpr.go.id/berita/detail/id/32146/t/Angka+Kekerasan+Meningkat%2C+RUU+PKS+Harus+Segera+Disahkan>.
- [3] A. Patros and C. Anggelia, "Polemik Desakan Pengesahan RUU PKS: Suatu Tinjauan Sistem," *Jurnal Komunikasi Hukum*, pp. 628-640, 2021.
- [4] A. T. Haryanto, "Pengguna Aktif Medsos RI 170 Juta, Bisa Main 3 Jam Sehari," 23 Februari 2021. [Online]. Available: https://kominfo.go.id/content/detail/2366/%20indonesia-peringkat-lima-penggunatwitter/0/sorotan_media.
- [5] B. Liu, "Sentiment Analysis and," May 2012. [Online]. Available: <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>. [Accessed 21 9 2021].
- [6] A. Go, R. Bhayani and L. Huang, "Twitter sentiment classification using distant supervision," *CS224N project report Stanford 1*, vol. no. 12, 2009.
- [7] A. Bermingham and F. A. Smeaton, "Classifying sentiment in microblogs: is brevity an advantage?," *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, pp. 1833-1836, 2010.
- [8] C. A. Melton, O. A. Olusanya, N. Ammar and A. Shaban-Nejad, "Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence," *Journal of Infection and Public Health*, no. JIPH-1670, p. 8, 2021.
- [9] A. S. Neogi, K. A. Garg, R. K. Mishra and Y. K. Dwivedi, "Sentiment analysis and classification of Indian farmers' protest using twitter data," *International Journal of Information Management Data Insights*, vol. 1, 2021.
- [10] K. Virra, R. Andreswari and M. A. Hasibuan, "Sentiment Analysis of Social Media Users Using Naïve Bayes, Decision Tree, Random Forest Algorithm: A Case Study of Draft Law on the Elimination of Sexual Violence (RUU PKS)," in *International Conference on Sustainable Engineering and Creative Computing (ICSECC)*, Bandung, 2019.
- [11] "MonkeyLearn," 2021. [Online]. Available: <https://monkeylearn.com/sentiment-analysis/>.
- [12] T. W. Edgar and D. O. Manz, "Chapter 4 - Exploratory Study," *Research Methods for Cyber Security*, pp. 95-130, 2017.
- [13] anand, "Medium," 29 8 2019. [Online]. Available: <https://medium.com/@atul.anand.sh/bernoulli-naive-bayes-over-svm-for-text-classification-77bdb5ac>. [Accessed 29 09 2021].
- [14] garydh, "garydh/paper-ep: v1," October 2021. [Online]. Available: <https://doi.org/10.5281/zenodo.5554534>.