Vol. 1 No. 1, March 2023 (11-20)
ISSN: 2987-6990
DOI: https://doi.org/10.60005/coreid.v1i1.4

Sentiment Analysis and Twitter Social Media Visualization Regarding the Omnibus Law Draft

Fadli Emsa Zamani
STISIP Samudera Indonesia Selatan

Article Info

Article history:

Received February 08, 2023
Revised February 24, 2023
Accepted March 21, 2023

Keywords:

Multi-Layer Perceptron
Omnibus law
Sentiment analysis
Visualization

ABSTRACT

This study will classify Twitter users' positive and negative opinions about the omnibus method using a frequency-inverse document frequency algorithm and a multi-layer perceptron method. The sentiment analysis process involves several stages, including B. Collecting and preprocessing data, calculating term weights using inverse term frequencies and document frequencies, and classifying data using multi-layer perceptrons. Additionally, the study visually represents Twitter's sentiment analysis results on the omnibus method. These visualizations include word cloud, top accounts, tweet frequency, hashtags, and sentiment. Three scenarios were considered to perform the classification experiments. Scenario 1 used 700 training data, scenario 2 used 800, and Scenario 3 used 900 training data. The findings show that the Term Frequency Inverse Document Frequency algorithm and the multi-layer perceptron method are adequate for sentiment analysis, with Scenario 3 yielding the highest accuracy rate of 88%.

Corresponding Author:

Fadli Emsa Zamani
STISIP Samudera Indonesia Selatan
Jl. Raya Pantai Cijayana, Cijayana, Kec. Mekarmukti, Kabupaten Garut, Jawa Barat 44165, Indonesia
Email: fadliemsa7@gmail.com

1. INTRODUCTION

Government is a form of organization that works with the task of running the government system. The government has the responsibility of distributing information, conveying policies, operational plans, and performance accomplishments to the public via multiple platforms, including traditional, conventional, and modern media[1]. The utilization of new media or internet technology for communication can efficiently and directly reach a wide range of individuals in a timely manner.

One way to get feedback on the public's response to the policy is to use social media. Based on data from the Association of Indonesian Internet Service Providers[2], the growth of internet users in 2020 in Indonesia has exceeded 143.26 million, which has an increase from the previous year in 2016-2018, namely the number of internet users reached 132.7 million and internet users ranging from 13-18 years of age are 16.68%, 19-34 years are 49.52%, 35-54 years are 29.55% and age over 54 years are 4.24%. The average internet user spends 1-3 hours per day, with an average number of users of 43.89%, and per week the average internet user is 65.98%[3].

The use of Twitter aside from being a medium for sharing information by posting various kinds of tweets, Twitter is also often used to socialize between users and express their sentiments or opinions on a topic or issue that is currently being discussed[4][5]. In the upcoming years, the fusion of data, information, and knowledge from various sources and organizations, along with integration, collaboration, and cooperation, as suggested by Gill Garcia [3], could become a significant research domain in the realm of government social media. In this study, the government's Twitter will be analyzed through sentiment analysis, which has not been widely explored before. One way to find out the public's response to implementing policies is to summarize opinions on social media. Social media contains information, opinions, and input from the public about many things[6]. A large number of Twitter users is often used by government agencies or entrepreneurs to monitor public opinion regarding ongoing government programs, including currently circulating issues that will be studied in this study, namely the draft omnibus law to obtain information. The information obtained from social networks can be in the form of positive or negative opinions. This Omnibus law has more to do with the government's work in the economic field[7] . The problem that most often becomes polemic is the omnibus law in the employment sector, the Job Creation Law. Similar to other legal languages, the term "omnibus" originates from the Latin word "omnis," meaning "many." Public opinions on social media are highly unstructured and do not conform to standardized sentence patterns. Thus, a technique is required to organize user opinions further to enable their analysis using a computational method called sentiment analysis.

The primary purpose of this study is to conduct sentiment analysis on Twitter and visualize the results to assess public sentiment towards omnibus law enforcement. This law was included in the government's national legislative program.

2. LITERATUR REVIEW

2.1. Omnibus law

The term "omnibus law" refers to regulations or laws covering a variety of subjects or issues. The literal meaning of the term is "law for all", derived from the Latin omnis, meaning "many" or "for all". According to Bryan A. Garner, the omnibus is defined in the 9th edition of the Black Law Dictionary as "referring to or dealing with many things at the same time. Containing many things or having various purposes.". This shows that the omnibus method concerns and addresses multiple objects and objects at the same time, serving different purposes. Known since 1840, this set of rules is a comprehensive set of rules not confined to a single set of rules. From this we can conclude that the omnibus law is a new law covering a wide range of regulatory issues and its existence amends several laws at once[2].

2.2. Sentiment Analisis

Sentiment analysis, also referred to as opinion mining, is a discipline that examines people's opinions, sentiments, evaluations, judgments, attitudes, and emotions towards entities such as products, services, organizations, individuals, problems, events, topics, and their attributes. Opinion mining can be applied to numerous domains/entities, including products and services, social and political events, and other activities. Opinions or viewpoints are the focal points of all human activity as they significantly impact our behavior. Sentiment analysis and opinion mining mainly focus on opinions that express or imply positive or negative sentiments. In early 2000, sentiment analysis had begun to develop into one of the active research areas in natural language processing (NLP). In recent years, sentiment analysis on industrial activity has also continued to grow[8].

The large number of studies on sentiment analysis is provided considerable benefits. One of its uses is that it can help to find out the opinion of the public or other people about a product by using it as a tool to find out the response of opinions about the product. Sentiment analysis, also called opinion mining, is a technique that examines people's opinions, ratings, judgments, attitudes, and feelings about various things such as products, services, organizations, individuals, issues, events, and their specific characteristics[9]. This field of study can be applied in various fields, including politics, to observe people's reactions to various issues. For example, emotional issues are illustrated in the case of a poll in which a person expresses her opinion about her Canon G12 camera. In this case, the statement consists of two components[10]

A positive, negative, or neutral target entity and mood. For example, sentences [10] and [11] express positive feelings about the Canon G12 camera and its image quality, while the sentence expresses negative feelings about the weight of the camera. According to Liu [12], the opinion definition includes purpose (G), mood (s), opinion creator (h), and time (t) at which the opinion was expressed.

2.3. Twitter API

Twitter was created by Jack Dorsey in 2006 and first launched in cyberspace in July 2006 with the address http://www.twitter.com which is still used today. Twitter API allows developers to access various features of the Twitter platform, such as retrieving tweets, searching for specific topics or hashtags, analyzing user data, and posting tweets on behalf of users. The API provides a set of rules, protocols, and tools that allow developers to interact with Twitter's servers and retrieve data in a structured format, such as JSON or XML. The documentation provides detailed information about the available endpoints, parameters, and authentication methods, as well as guidelines and best practices for using the API[13]. There are several types of Twitter APIs such as:

2.3.1. Twitter REST API

Consists of Twitter REST and Twitter Search. Twitter REST provides core data and core Twitter objects. Twitter search also functions to search for an instance of a Twitter object or search for trends[14].

2.3.2. Twitter Streaming API

The Twitter API is commonly used for data mining because it allows developers to access a large amount of real-time data from Twitter. This data can be used for various purposes, such as sentiment analysis, trend analysis, and social network analysis. With the Twitter API, developers can retrieve tweets and other information from Twitter, such as user profiles, hashtags, and mentions, and then analyze this data using various techniques and tools [15].

2.4. Multi-layer Perceptron

Multi-layer Perceptron (MLP) is an ANN derived from Perceptron, in the form of a feedforward ANN with one or more hidden layers. In a neural network, the input layer receives the input signal, which is then processed by the hidden layer(s) and finally, the output layer produces the output signal[16]. The input signal is passed forward through the layers in a process called forward propagation, which involves the computation of weighted sums and activation functions in each neuron. The output signal is then compared to the expected output, and the error is propagated backward through the network in a process called backpropagation, which updates the weights and biases of the neurons to minimize the error[17][18]. An example of an MLP architecture is given below:

Commercial artificial neural networks (ANNs) usually comprise three or four layers, including one or two hidden layers. These layers can contain anywhere from 10 to 100 neurons. The input signal is propagated in a forward direction through the layers. The design and architecture of the ANN are important factors that can affect its performance in terms of accuracy and speed of processing. Experimental ANNs can have five or even six layers (three or four hidden layers) where each layer uses millions of neurons, but in most applications, it uses two layers (one more hidden layer) because each addition of one layer will increase the computational load exponentially [19].

The following is the Multi-layer Perceptron algorithm [20]:

a. Begin by initializing all weights with small random values.
b. If the termination conditions have not been met, move to steps 2-8.
c. Perform steps 3-8 for each pair of training data.
d. Send a signal to each input unit, which then passes it on to the corresponding hidden unit above it.
e. Compute all the outputs in the hidden unit $z\ j$ (j = 1, 2,..., p).

$$z\_net_j = v_{j0} + \sum_{i=1}^{n} x_i\, v_{ji} \tag{1}$$

f. Count all network outputs in output units $y\ k$ (k = 1, 2,...,m).
g. Calculate the δ factor of output units based on the error in each output unit yk (k = 1, 2,..., m).

$$\delta k = (tk - yk)'(y\_netk) = (tk - yk)yk(1 - yk), tk = target \tag{2}$$

δk is the error unit that is used to adjust the weights of the layer below it. The change in weight wkj is calculated by multiplying δk with the learning rate α.

$$\Delta wkj = \alpha\delta\ kZj,\ k=1,2\ \ldots\ .,; j=0,1\ldots p \tag{3}$$

h. Calculate the hidden unit δ factor based on the error in each hidden unit zj (j = 1).

$$\delta\_net_j = \sum_{k=1}^{m} \delta_k w_{kj} \tag{4}$$

Hidden unit δ factor.

$$\delta j = \delta\_netjf'(Z\_netj) = \delta\_net\ jzj(1 - zj) \tag{5}$$

Calculate the weight change rate vji

$$\Delta vji = \alpha\delta\ jx\ ,\ j=1,2,\ \ldots\ ,p\ ;1=1,2,\ \ldots\ .,n \tag{6}$$

i. Count all the weight changes, including the changes in weights of the lines leading to the output METODEunit, which can be calculated using the output error δo and the activation value of the hidden unit:

$$wkj(\text{new}) = w\ kj(\text{old}) + \Delta wkj,(k=1,2,\ \ldots\ ,m\ ;j=0,1\ldots\ .,p) \tag{7}$$

The change in weight of the line going to the hidden unit, ie:

$$Vji(\text{new}) = v\ ji(\text{old}) + \Delta vji\ ,(j=1,2,\ \ldots\ .,p;i=0,1,\ \ldots\ .,n) \tag{8}$$

3. METHOD

3.1. Input

The input entered by the system is a document in the form of a tweet from a Twitter account in the form of an opinion. The tweet data is obtained by utilizing the API (Application Interface) feature that has been provided by Twitter. The documents submitted are Indonesian language documents.

3.2. Datasets

The dataset is in Indonesian text format downloaded from the Twitter website. The data collected for this study was obtained using the omnibus method, job creation, worker, chilla, and creator queries. These queries are numerous accounts such as B. Accounts @jokowi, and @kemenkeu. The tweets retrieved are the posts from the query. A flowchart of the record retrieval process. The dataset from the results of this crawl is split into two parts, training data, and test data, as shown in Table 7. This crawled data is in the form of tweet documents with no other attributes. This training data is entered into a database and manually categorized with positive or negative mood labels. Storing the test data obtained from the Twitter API crawling process in a database is a necessary step for further processing by the system[21]. This processing allows the system to automatically generate output in the form of positive or negative sentiment.
The preprocessing process is important in the next stage which is attribute reduction which has less impact on the classification process. Since the data entered at this stage is still raw and dirty data, the result of this process is a high-quality document that is expected to facilitate the classification process[22]. The pretreatment process consists of several stages.

3.3. Case Folding

In the case folding stage, all uppercase letters in the tweet documents are converted to lowercase letters. The purpose of this stage is to remove redundant data that only differ in their letter cases. As an illustration of the case folding process, the following examples of the resulting tweets are shown in Table 1.

Table 1. Case Folding Sample Fase

Input Process	Output Process
@imamarina3 Dengan adanya Omnibus Law, Buruh mendapatkan Bonus 5 kali Gaji #tolakdemoaksisaatpandemi	@imamarina3 dengan adanya omnibus law, buruh mendapatkan bonus 5 kali gaji #tolakdemoaksisaatpandemi

3.2. Cleansing

During the cleansing stage, irrelevant components in the tweeted document are removed, which do not affect the sentiment classification results. These components include mentions starting with the attribute '@', hashtags starting with the attribute '#', and links starting with the attribute 'http' or 'bit.ly', as well as symbol characters such as '~!@#$%^&*()_+?<>,.?:{}[]|'. They are replaced with a space character to maintain the structure of the tweeted document.
As an illustration of the following cleaning process, the author provides an example of the resulting tweets as shown in Table 2.

Table 2. Cleansing Fase Sample

Input Process	Output Process
@imamarina3 dengan adanya omnibus law, buruh mendapatkan bonus 5 kali gaji #tolakdemoaksisaatpandemi	dengan adanya omnibus law, buruh mendapatkan bonus 5 kali gaji #tolakdemoaksisaatpandemi

3.3. Tokenizing

The tokenizing stage involves breaking down the words in a document into individual pieces. In the case of tweets, words are separated by spaces. This process results in individual words that are entered into the database for weighting purposes[23]. The following is a tokenizing flow chart as an illustration of the tokenizing process along with examples of tweets generated as shown in Table 3.

Table 3. Tokenizing sample fase

Input Process	Output Process
pesanan sudah dalam pengiriman, jadi sudah tidak bisa dibatalkan	‘pesanan’ ‘sudah’ ‘dalam’ ‘pengiriman’ ‘jadi’ ‘sudah’ ‘tidak’ ‘bisa’ ‘dibatalkan’

3.4. Stopword Removal

The stopword removal stage involves removing words that are not relevant to the topic of the document, as these words do not affect the accuracy of sentiment classification[24]. A stopword list is created containing such words, and if a word in the tweeted document matches a word in the stopword list, it will be removed and replaced with a space character. The following is a stopword removal flowchart. As an illustration of the stopword removal process, the following is an example of a tweet that has been carried out by the stopword removal process, which will produce documents as shown in Table 4.

Table 4. Stopword Removal Sampel Fase

Input Process	Output Process
bukan kah resinya sudah saya kirim dan juga sudah di konfirm oleh lazada klo barangnya sudah sampai di gudang lazada	bukan resinya kirim konfirm lazada klo barangnya gudang lazada

3.5. Stemming

The process of stemming involves applying specific rules to change the words in a document into their root form. In Indonesian stemming, this is achieved by removing any suffixes, prefixes, and confixes found in the text. The method used in this process was developed by Bobby Nazief and Mirna Adriani and involves several stages[25]. First, the word is looked up in a dictionary and if found, it is assumed to be the root word. If not, inflection suffixes such as "-lah", "-kah", "-ku", "-mu", or "-nya" are removed, along with any possessive pronouns if present. Next, derivation suffixes like "-i", "-an", or "-kan" are removed and if the word is found in the dictionary, the method stops. Otherwise, additional steps are taken to try to find the root word. If all steps have been completed but the root word has not been found, the original word is assumed to be the root word. To make the stemming process perfect, you have to go through several stages of the processes mentioned in the points above. As an illustration of the stemming process, the following is an example of the resulting tweet as shown in Table 5.

Table 5. Table Stemming Procces

Input Process	Output Process
Undang undang cipta kerja akan memebrikan kemudahan untuk para pekerja	Undang undang cipta kerja akan beri mudah untuk para kerja

4. RESULTS

4.1. Classify using Multi-layer Perceptron

Classification is done using the multi-layer perceptron (MLP) algorithm by using the weight of each term selected from the tweets that have been obtained. The value used is the TF-IDF value of each word that has the highest weight.

Table 6. Multi-layer Perceptron Parameter's

No	Parameter	Information
1	Input layer	300 nodes
2	Hidden layer 1	117 nodes
3	Hidden layer 2	117 nodes
4	Output layer	1 node
5	Activation function	Binary sigmoid
6	Learning rate	0.05

4.2. Calculate Multi-layer Perceptron

The following is a manual calculation of the multi-layer perceptron (MLP) method. If the tf-idf value and label are known as follows:

Initialize all weights with small random numbers. This initialization is done by giving a random value on each side that connects between nodes. Then calculate the values of all input nodes multiplied by the side that connects between nodes to get the value that will be entered into the sigmoid activation function, this process is referred to as feedforward.

Table 7. Trial Scenario

Scenario	Number of training data		Number of test data
	Positive	Negative	
1	350	350	100
2	400	400	100
3	450	450	100

Here are some sample tweets from each scenario as follows:

Table 8. Tweet each Scenario

Scenario	*Tweet* Positive	Negative
1	kinerja Positive pemerintah jokowi dalam membangun infrastruktur dan meciptakan lapangan pekerjaan dengan omnibus law	antek aseng kriminalisasi ulama..saya yang mendzolimi atau ada yang mendzolimi ke saya ~ Jokowi
	Cuma di era Jokowi anggaran kedaulatan pangan naik	Berbagai isu fitnah...antek asing aseng PKI gak laku!!! Malah bikin elektabilitas Jokowi semakin meroket
2	Percayalah mayoritas rakyat sudah pintar tahu mana yang kerja bener dan hasilnya benar	Salah satunya didemo karyawan kertas nusantara karena gak bayar gaji dan THR
	Cipta kerja harus segera dilaksanakan pemerintah	Omnibus law hanya menguntungkan pengusaha
3	Investor akan datang jika regulasi dipermudah dengan omnibus law	Yang merancang RUU cita kerja ini sebenarnya DPR tapi tetap saja presiden yang mengesahkan
	Sebetulnya omnibus law ini tujuannya baik untuk menciptakan lapangan pekerjaan yang banyak	Jokowi dan pdip bersekongkol untuk mengesahkan ruu omnibus law untuk kepentingan china

4.3. Result of Classification of the sentiment using TF-IDF Algorithm and MLP

The sub-chapter presents the classification results of sentiment analysis performed on public opinion regarding the omnibus law in Indonesia. The classification was conducted using the TF-IDF algorithm and the multi-layer perceptron (MLP) neural network method. The accuracy level of the algorithm was also determined. The output of the sentiment classification research was presented in the form of a table and graphs obtained from the scenario testing process on the system developed by the author using the Python programming language. The table displays the results of the classification output for each scenario:

4.3.1. Scenario 1 Accuracy Level Results

Table 9. Scenario 1 Accuracy Level Results

Sentiment		Prediction result class Positive	Negative
Actual class	Positive	a = 34	b = 7
	Negative	d = 8	c = 50

$$\text{Accuracy} = \frac{34+50}{50+7+10+34} \text{ x } 100$$

$$\text{Accuracy} = \frac{83}{100} \text{ x } 100 = 83\ \%$$

4.3.2. Second Scenario Accuracy Level Results

Table 10. Scenario Accuracy Level 2 Result

Sentiment		Prediction result class Positive	Negatives
Actual class	Positive	a = 32	b = 7
	Negative	d = 8	c = 52

Based on the test results from the table above, the accuracy value can be taken as follows:

$$\text{Accuracy} = \frac{52+32}{52+7+8+32} \text{ x } 100$$

$$\text{Accuracy} = \frac{85}{100} \text{ x } 100 = 85\ \%$$

4.3.3. Scenario 3 Accuracy Level Results

Table 11. Scenario Testing Results 3

Sentiment		Prediction result class	
		Positive	**Negative**
Actual class	Positive	a = 30	b = 7
	Negative	d = 5	c = 54

Based on the test results from the table above, the accuracy value can be taken as follows:

$$accuracy = \frac{30+54}{30+7+5+54} \text{ x } 100$$

$$accuracy = \frac{88}{100} \text{ x } 100 = 88\ \%$$

The variables in the classification process are represented by four parameters: a, b, c, and d. The parameter 'a' represents the number of positive records correctly classified as positive, 'b' represents the number of positive records incorrectly classified as negative, and 'c' represents the number of negative records correctly classified as negative. , and 'd' represents the number. Negative records are misclassified as positive. Overall, the test results for the three scenarios in this study are summarized based on the following graphs. The results of the accuracy obtained from the three scenarios above show that the trend of discussion around the omnibus law in the discussion of employment shows negative sentiment from Twitter users.

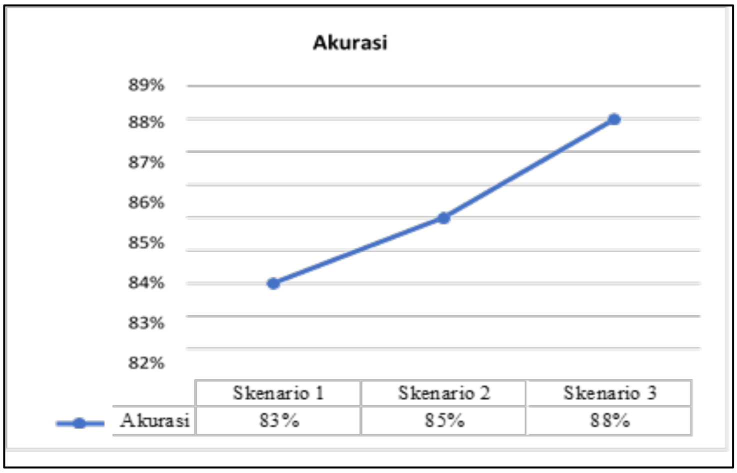

Figure 1. Accuracy of three schema

4.4. Sentiment Analysis Visualization

4.4.1. Omnibus Law Trends

Dandhy Laksono is the person whose tweet gets the most attention on this topic because.

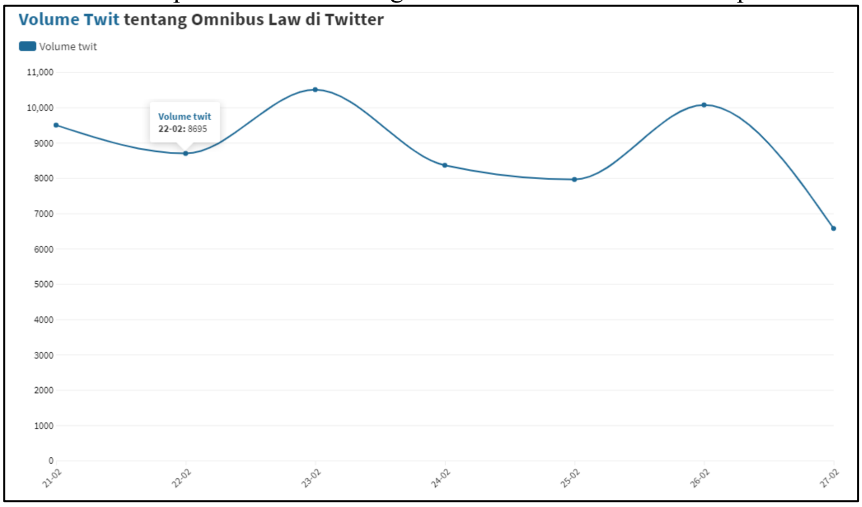

Figure 2. Volume Tweet about omnibus law topic

Meanwhile, Tempo is the media that gets the most attention. The following is a list of top accounts hat often discuss opinions about the omnibus law which then receive many responses from other Twitter users, the number of which is calculated based on the number of likes and retweets received.

Table 12. Top Account

No.	*Account Name*	*RT + Reply Count*
1	@Dandhy_Laksono	12,260
2	@do_ra_dong	5,266
3	@xaliber	3,065
4	@MichelAdamNP	2,722
5	@ustadtengkuzul	2,699
6	@geloraco	2,590
7	@podoradong	2,449
8	@tempodotco	2,413
9	@TirtoID	2,177
10	@K1r4n14__	2,118
11	@CNNIndonesia	2,090

4.4.2. Sentiment Visualization

The majority of Twitter users gave negative responses to the Omnibus Law Bill. Even so, the positive response of 32% on this topic is also quite large.

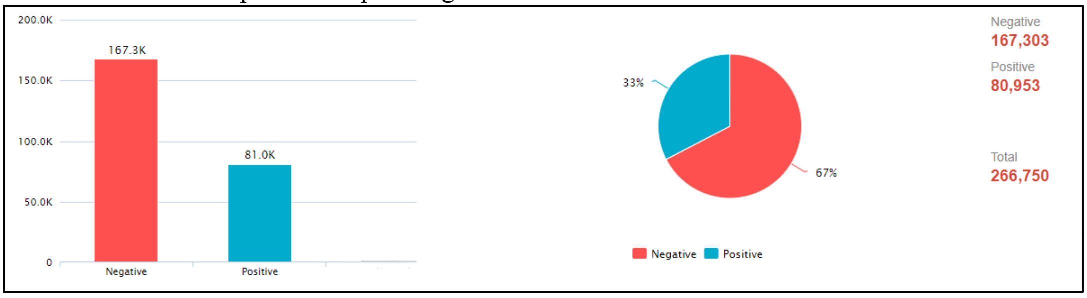

Figure 3. Twitter user sentimen

4.4.3. Emotion

Angry and shocked emotions describe most of the feelings of netizens. The rest responded to this bill with feelings of sadness, wariness, and a feeling of trust, and disbelief in the purpose of drafting this law.

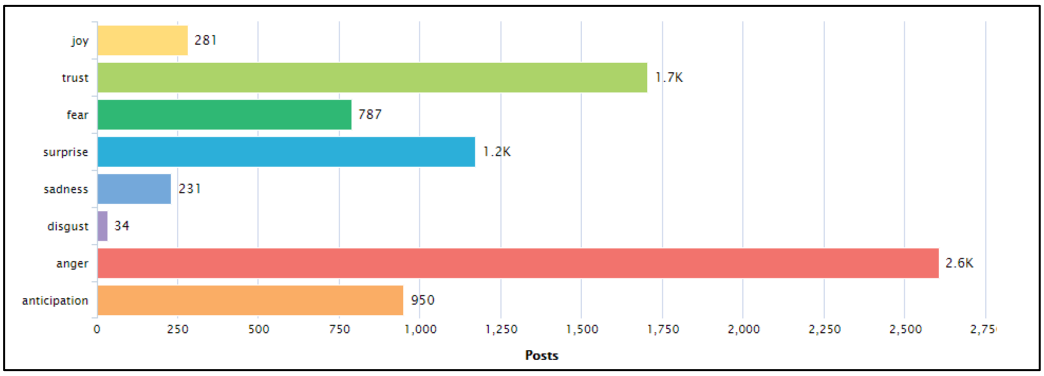

Figure 4. User Emotions on the Omnibus Law topic

4.4.4. Social Network Analysis

There are two clusters that show the pros and cons in discussing the omnibus law. The top influencers on this topic are also clearly visible in the SNA. In addition, even though they are not taking sides, the media that report this topic negatively are also included in the red cluster.

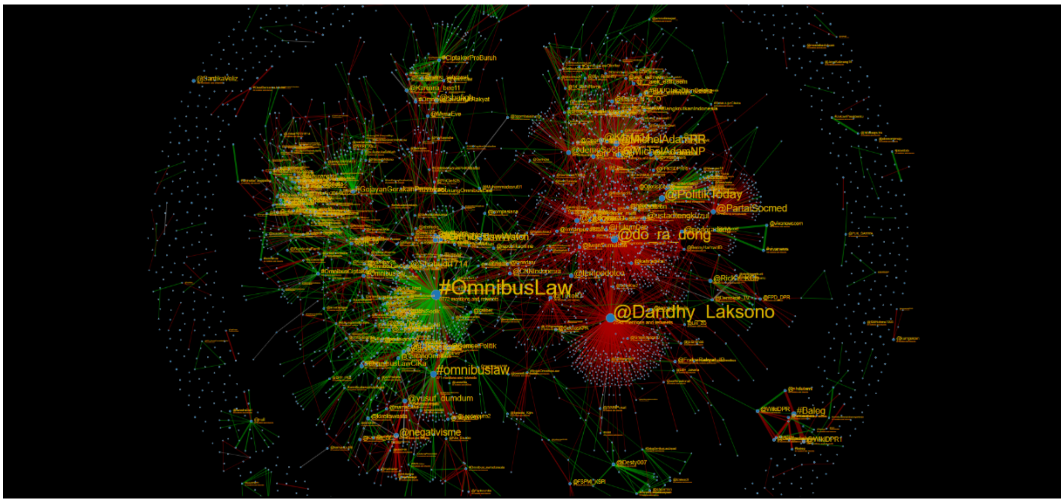

Figure 5. Social Media Node Analysis

4.4.5. Frequency of Tweet

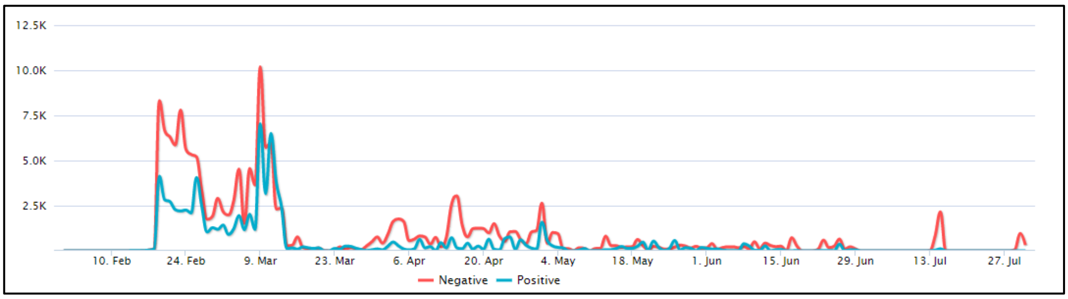

Figure 6. Frequency of Tweet

The following is the frequency of tweet data for Twitter users discussing the topic regarding this omnibus law, it can be seen from the graph that this topic was widely discussed in February, there were many pros and cons from the beginning of this draft law, up to the July data the omnibus law topic was still ongoingly discussed.

5. CONCLUSION

Summarizing the previous chapter, we described the successful implementation of the term frequency-inverse document frequency (TF-IDF) algorithm and the multi-layer perceptron (MLP) neural network method for Twitter sentiment analysis. The system was able to provide both positive and negative sentiment output values, with Scenario 3 achieving the highest accuracy value of 88%. From this, we can conclude that these methods of sentiment analysis on Twitter are valid.

REFERENCES

[1] H. K. Duan, M. A. Vasarhelyi, M. Codesso, and Z. Alzamil, "Enhancing the government accounting information systems using social media information: An application of text mining and machine learning," *Int. J. Account. Inf. Syst.*, vol. 48, p. 100600, Mar. 2023, doi: 10.1016/J.ACCINF.2022.100600.

[2] S. Al-Fatih, "Harmonisasi Regulasi Hubungan Pusat & Daerah Melalui Omnibus Law (Harmonizing Regulation Of Central & Local Regulations Through The Omnibus Law)," *SSRN Electron. J.*, Jul. 2020, doi: 10.2139/SSRN.3857493.

[3] B. Pang and L. Lee, "Opinion mining and sentiment analysis," 2008.

[4] S. Bayrakdar, I. Yucedag, M. Simsek, and I. A. Dogru, "Semantic analysis on social networks: A survey," *Int. J. Commun. Syst.*, vol. 33, no. 11, p. e4424, Jul. 2020, doi: 10.1002/DAC.4424.

[5] U. Can and B. Alatas, "A new direction in social network analysis: Online social network analysis problems and applications," *Phys. A Stat. Mech. its Appl.*, vol. 535, p. 122372, Dec. 2019, doi: 10.1016/J.PHYSA.2019.122372.

[6] L. Hudders, S. De Jans, and M. De Veirman, "The commercialization of social media stars: a literature review and conceptual framework on the strategic use of social media influencers,"

https://doi.org/10.1080/02650487.2020.1836925, vol. 40, no. 3, pp. 327–375, 2020, doi: 10.1080/02650487.2020.1836925.

[7] T. Okumura and Y. Kanatani, "Health Crisis Management and Natural Language Processing," *J. Nat. Lang. Process.*, vol. 20, no. 3, pp. 513–524, 2013, doi: 10.5715/JNLP.20.513.

[8] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: state of the art, current trends and challenges," *Multimed. Tools Appl.*, vol. 82, no. 3, pp. 3713–3744, Jan. 2023, doi: 10.1007/S11042-022-13428-4/FIGURES/3.

[9] F. Hemmatian and M. K. Sohrabi, "A survey on classification techniques for opinion mining and sentiment analysis," *Artif. Intell. Rev.*, vol. 52, no. 3, pp. 1495–1545, Oct. 2019, doi: 10.1007/S10462-017-9599-6/METRICS.

[10] E. Dwianto and M. Sadikin, "Analisis Sentimen Transportasi Online pada Twitter Menggunakan Metode Klasifikasi Naïve Bayes dan Support Vector Machine," *Format J. Ilm. Tek. Inform.*, vol. 10, no. 1, pp. 94–100, Feb. 2021, doi: 10.22441/FORMAT.2021.V10.I1.009.

[11] R. Ferdiana, F. Jatmiko, D. D. Purwanti, A. Sekar, T. Ayu, and W. F. Dicka, "Dataset Indonesia untuk Analisis Sentimen," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 8, no. 4, pp. 334–339, Nov. 2019, Accessed: May 20, 2023. [Online]. Available: https://jurnal.ugm.ac.id/v3/JNTETI/article/view/2558

[12] B. Lu, M. Ott, C. Cardie, and B. Tsou, *Multi aspect sentiment analysis with topic models*, IEEE 11 th. 2011.

[13] N. Thakur, "Twitter Big Data as a Resource for Exoskeleton Research: A Large-Scale Dataset of about 140,000 Tweets from 2017-2022 and 100 Research Questions," *Anal. 2022, Vol. 1, Pages 72-97*, vol. 1, no. 2, pp. 72–97, Sep. 2022, doi: 10.3390/ANALYTICS1020007.

[14] C. Udanor and C. C. Anyanwu, "Combating the challenges of social media hate speech in a polarized society: A Twitter ego lexalytics approach," *Data Technol. Appl.*, vol. 53, no. 4, pp. 501–527, Oct. 2019, doi: 10.1108/DTA-01-2019-0007/FULL/XML.

[15] S. F. Rodiyansyah and E. Winarko, "Klasifikasi Posting Twitter Kemacetan Lalu Lintas Kota Bandung Menggunakan Naive Bayesian Classification," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 6, no. 1, Jan. 2013, doi: 10.22146/ijccs.2144.

[16] Z. Kalinić, V. Marinković, L. Kalinić, and F. Liébana-Cabanillas, "Neural network modeling of consumer satisfaction in mobile commerce: An empirical analysis," *Expert Syst. Appl.*, vol. 175, p. 114803, Aug. 2021, doi: 10.1016/J.ESWA.2021.114803.

[17] A. H. S. Hamdany, R. R. O. Al-Nima, and L. H. Albak, "Translating cuneiform symbols using artificial neural network," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 19, no. 2, pp. 438–443, Apr. 2021, doi: 10.12928/TELKOMNIKA.V19I2.16134.

[18] T. C. My, L. D. Khanh, and P. M. Thao, "An Artificial Neural Networks (ANN) Approach for 3 Degrees of Freedom Motion Controlling," *JOIV Int. J. Informatics Vis.*, vol. 7, no. 2, pp. 301–309, May 2023, doi: 10.30630/JOIV.7.2.1817.

[19] M. S. Al-Batah, S. Mrayyen, and M. Alzaqebah, "Arabic Sentiment Classification using MLP Network Hybrid with Naive Bayes Algorithm," *J. Comput. Sci.*, vol. 14, no. 8, pp. 1104–1114, Aug. 2018, doi: 10.3844/JCSSP.2018.1104.1114.

[20] R. Melita, V. Amrizal, H. B. Suseno, and T. Dirjam, "Penerapan Metode Term Frequency Inverse Document Frequency (TF-DF) dan Cosine Similarity pada Sistem Temu Kembali Informasi untuk Mengetahui Syarah Hadits Berbasis Web (Studi Kasus: Hadits Shahih Bukhari-Muslim)," *J. Tek. Inform.*, vol. 11, no. 2, pp. 149–164, Nov. 2018, doi: 10.15408/JTI.V11I2.8623.

[21] A. R. Atmadja and A. Purwarianti, "Comparison on the rule based method and statistical based method on emotion classification for Indonesian Twitter text," in *2015 International Conference on Information Technology Systems and Innovation, ICITSI 2015 - Proceedings*, 2016. doi: 10.1109/ICITSI.2015.7437692.

[22] M. Sumathi and S. A. Parvin, "Nuances of data pre-processing and its impact on business," *Proc. - 5th Int. Conf. Intell. Comput. Control Syst. ICICCS 2021*, pp. 1006–1012, May 2021, doi: 10.1109/ICICCS51141.2021.9432376.

[23] M. Alsolamy, A. M. Alotaibi, A. Alabbas, M. Abdullah, M. Abdulaziz, and A. Alotaibi, "Topic based Sentiment Analysis for COVID-19 Tweets," *Artic. Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 1, p. 2021, 2021, doi: 10.14569/IJACSA.2021.0120172.

[24] A. K. Uysal and S. Gunal, "The impact of preprocessing on text classification," *Inf. Process. Manag.*, vol. 50, no. 1, pp. 104–112, Jan. 2014, doi: 10.1016/J.IPM.2013.08.006.

[25] J. Asian, H. E. Williams, and S. M. M. Tahaghoghi, "Stemming Indonesian," *ACM Trans. Asian Lang. Inf. Process.*, vol. 38, pp. 307–314, Dec. 2007, doi: 10.1145/1316457.1316459.