

Detecting Cyberbullying on Social Media Using Support Vector Machine: A Case Study on Twitter



Al-Khowarizmi^{1*}, Indah Purnama Sari¹, Halim Maulana²

¹ Department of Information Technology, Universitas Muhammadiyah Sumatera Utara, Medan 20238, Indonesia

² Department of Information System, Universitas Muhammadiyah Sumatera Utara, Medan 20238, Indonesia

Corresponding Author Email: alkhowarizmi@umsu.ac.id

<https://doi.org/10.18280/ijssse.130413>

ABSTRACT

Received: 2 July 2023

Revised: 15 August 2023

Accepted: 20 August 2023

Available online: 28 September 2023

Keywords:

support vector machine (SVM), detection, cyberbullying, social media, kernel function

Cyberbullying, a prevalent issue in digital media, particularly social media, poses a significant concern owing to its pervasive nature and potential harm. Social media platforms permit users to exchange opinions freely, which, while fostering open discourse, can also trigger instances of cyberbullying. This study focusses on Twitter discourses related to Indonesia's contentious public policy, "Cipta Kerja". The inherent polarity of views towards this policy has given rise to instances of cyberbullying. An extensive dataset comprising 2400 tweets was meticulously assembled, employing the keyword "Cipta Kerja". This dataset was subsequently partitioned into training and testing subsets to facilitate cyberbullying detection through computational algorithms. Sentiment analysis played a crucial role in this process, with the Support Vector Machine (SVM) method demonstrating remarkable reliability in classifying sentiment-related issues and, therefore, detecting cyberbullyings. The SVM method, using a linear kernel function, achieved a commendable accuracy rate of 92.7% in cyberbullying detection. This study's results underscore the effectiveness of SVM in identifying instances of cyberbullying on social media platforms, offering new promise for safeguarding digital spaces.

1. INTRODUCTION

Cyberbullying, a form of bullying enacted through digital technology, is a disturbing trend that manifests in cyberspace, often resulting in humiliation and intimidation [1, 2]. The repercussions of cyberbullying are primarily socio-psychological, exerting significant pressure on its victims [3]. It is commonly observed on digital platforms like social media and online gaming environments [4]. A 2022 UNICEF report revealed that out of 2,777 child respondents from Indonesia, an alarming 45% claimed to have been victims of cyberbullying [5]. Motivations for individuals to engage in cyberbullying often include the pursuit of popularity, a lack of awareness about its impact, or a need to assert dominance [6]. In light of these findings, the role of parental supervision has been emphasised in combating cyberbullying, especially in regulating children's internet use on smartphones and their access to social media and specific content [7]. Recent surveys by the Association of Indonesian Internet Service Providers (APJII) predict that by 2023, internet users will reach up to 215.63 million, with the highest frequency of usage among adolescents aged 15-20 years [8, 9]. On average, these teenagers utilize the internet for accessing social media platforms and consuming various content [10]. The number of social media users in Indonesia hit 277.7 million in 2022, a figure that continues to rise with advancing technology [10].

Social media platforms are recognized for their multifaceted advantages, including facilitating communication among individuals, enabling information exchange, and fostering business development [11]. Twitter, one of the popular social media platforms, has seen a significant increase in active users

and has become a widely used platform for information exchange [12]. Despite limiting posts or 'tweets' to 280 characters, Twitter effectively facilitates communication and opinion sharing on social and political topics [13, 14].

The user-generated data from tweets contain valuable information that can be leveraged using data mining techniques, thereby aiding decision-makers [15, 16]. Text mining, a subset of data mining, is employed on communication data between individuals on Twitter with the objective of generating information related to sentiment analysis [17, 18]. This process involves extracting useful information from tweet data to identify sentiment patterns within unstructured data [17, 18]. Sentiment analysis, the process of analyzing opinions from a text, is aimed at identifying whether the data contains positive, negative, or neutral sentiments [19, 20]. Data derived from posts on Twitter will be classified into three categories: positive sentiment, negative sentiment, and neutral sentiment [19]. This classification involves several stages including feature extraction, text preprocessing, and evaluation.

The Support Vector Machine (SVM) is an algorithm known for its robust performance in classification tasks [21]. As a kernel-based machine learning model, SVM has demonstrated potential in extracting information with comparatively smaller training data sets [22]. The SVM algorithm is versatile and can process both linear and non-linear data. In the case of linear data, SVM provides an efficient dimensionality for small data sets. For non-linear data, the algorithm employs kernel functions [23, 24]. Previous research emphasizes the efficacy of SVM in sentiment analysis. A study conducted by Hamdi [25] integrated an ant colony optimization algorithm with

SVM for sentiment classification using review data from Amazon, achieving an accuracy rate of 89%. Another study by Ahsaan et al. [26] combined a linear SVM with a large dataset from social media and IoT, resulting in an accuracy of 91% in sentiment classification. The current study focuses on Twitter post data, characterized by its non-linear properties due to the complex nature of the content, which often includes numerous abbreviations, hashtags, emoticons, and punctuation. This complexity renders Twitter post data as unstructured. Consequently, the kernel function of SVM will be employed in this study. The linear kernel is one of the kernel functions in SVM and has been chosen for this study due to its ability to solve sentiment classification problems effectively. The linear kernel processes data into a broader dimensional feature space without necessitating transformation in high dimensions [27]. The purpose of using a linear kernel function in this paper is to optimize kernel performance for more effective results. The linear kernel is particularly useful for understanding the relationships between features and classes.

2. MATERIAL AND METHOD

2.1 Data and general research architecture

The data in this study uses data on Twitter that carries out bullying so that it is classified as cyberbullying. The comments and tweets that are grouped are related to public policy. Various public policies are of concern to many Twitter users and become the pros and cons associated with these public policies. Both the pros and cons give rise to bullying so it is necessary to detect cyberbullying. So that the general architecture in this paper is formed as shown in Figure 1.

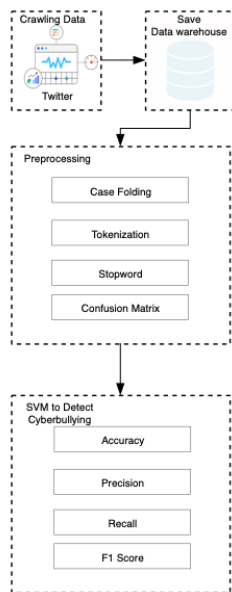


Figure 1. General research architecture

Figure 1 is the general architecture in this study. The use of general architecture so that this research is in line and in accordance with the results to be aimed at. There is also a step of the general architecture as follows:

1. Crawling from from twitter with the keyword "Cipta Kerja" in Indonesian
2. Perform data storage on data driven
3. Doing preprocessing

- a. Do case folding
 - b. Tokenization
 - c. Make a stop word on a meaningful word
 - d. Perform data visualization using a confusion matrix
4. Detect with SVM and get:
 - a. Accuracy based on Eq. (1) [27]

$$Accuracy = \frac{TP + TN + TN}{(TP + TN + FP + FN + TN + FP + FN + TN + TP)} \times 100\% \quad (1)$$

- b. Precision based on Eq. (2) [28]

$$Precision = \frac{TP}{(TP + FP)} \quad (2)$$

- c. Recall with Eq. (3) [29]

$$Recall = \frac{TP}{(TP + FN)} \quad (3)$$

- d. F1-Score based on Eq. (4) [30]

$$F1 - Score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (4)$$

2.2 Support vector machine

SVM is a very popular method in machine learning for solving classification and regression problems [31]. The main goal is to separate the two data classes using a hyperplane with the maximum distance from the data points closest to the dividing line. In SVM, data is represented as points in a multidimensional space, where each dimension represents features of the data [32]. SVM tries to find a hyperplane that maximizes the margin and distance between two classes of data. In addition, SVM can also use kernel functions to transform the feature space into a higher dimensional space, which allows for better separation if the data is not linearly separable. SVM has several advantages, including the ability to work well on high-dimensional datasets, the ability to overcome overfitting problems, and is effective in cases where class separation is not linear [33]. However, SVM also has the disadvantage of slow performance for very large datasets. Basically SVM is measured by hyperline based on Eq. (5) [34].

$$W \cdot x \phi(x) + b = 0 \quad (5)$$

where, W is the weight value of the hyperline, $\phi(x)$ is a function of the input vector, and b is the bias value. Then from these calculations an optimization is carried out in order to get the minimum value based on Eq. (6) [35]:

$$\min_{\alpha} \psi(\alpha) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j K(x_i, x_j) \alpha_i \alpha_j - \sum_{i=1}^N \alpha_i \quad (6)$$

where, x_i is the input vector value, y_i is the n value of the class correspondence, α is the Lagrange multiplier value, and K is a linear kernel function. In the combination of parameters must be determined in advance in getting it. Thus, SVM can be grouped into linear and non-linear. As seen in Eq. (7) [36]:

$$(x) = \sum_{i=0}^N \alpha_i y_i x_i^T \cdot x + \beta_0 \quad (7)$$

where, x_i is the value on the labelled of y_i , α is the value to be multiplied by Lagrange and β_0 is the bias value of 0. In non-linear problems the equation is changed based on Eq. (8) [28].

$$f(x) = \sum_{i=0}^N \alpha_i y_i K(x_i, x) \cdot x + \beta_0 \quad (8)$$

where, the value of N is the value of the number of input vectors. So that an algorithm is formed to detect cyberbullying with SVM based on the following pseudocode.

Pseudocode 1: SVM Algorithms for Cyberbullying Detection

Input: Training dataset (X, y), where X is the feature matrix and y is the label vector

C: regularization parameter

kernel: type of kernel function used linear)

max_iterations: maximum number of iterations

Output: Trained SVM model

1. Initialize alpha_i, b, and iteration = 0

2. Define the kernel function:

- linear_kernel(X_i, X_j):

return dot_product(X_i, X_j)

3. Define the decision_function(X) function:

- result = 0

- for i in range(m):

- if alpha_i > 0:

- result += alpha_i * y_i * kernel(X_i, X)

- return result + b

4. Define the train_svm() function:

- while iteration < max_iterations:

- iteration += 1

- for i in range(m):

- compute the decision function for sample i:

- F_i = decision_function(X_i)

- compute the error value for sample i:

- error_i = y_i * F_i - 1

- update alpha_i based on the rule:

- alpha_i = alpha_i - (error_i / (||X_i||^2 + 1 / (2C)))

- bound alpha_i between 0 and C:

- alpha_i = max(0, min(alpha_i, C))

- compute the bias term b based on the mean of errors for samples that have alpha_i > 0:

- b = mean(y_i - decision_function(X_i) for alpha_i > 0)

5. Call the train_svm() function to train the SVM model.

6. Output the trained SVM model.

3. RESULT AND DISCUSSION

In detecting cyberbullying, of course, we go through the sentiment analysis stage of the data provided. The research contained in this paper is of course data crawling results from Twitter social media based on public policies that are informed on social media so that the wider community is aware of these policies. From the public policies circulating on Twitter, of course, it raises the pros and cons of a democratic society so that the result of the message becomes bullying in the digital media. In getting accurate results of course the data crawled as many as 2400 tweets with the keyword "Cipta Kerja" in Indonesian. The data is converted into structural data by preprocessing and labelling. The results are shown in Table 1.

Table 1. Summary of preprocessing

Tweet	Result of case folding	Tokenization Results	Stopword results
@cholilnafis Hopefully the Constitutional Court's decision will not be like the fate of the Cipta Kerja Law. Kicked using the ultimate move called fake Perpu...	Hopefully this decision of MK will not be like the fate of those who create jobs being kicked out using the ultimate trick called a fake perpu.	'hopefully', 'decision', 'mk', 'this', 'don't', 'like', 'fate', 'uu', 'create', 'work', 'in', 'kick', 'use', 'stance', 'ultimate', 'yang', 'name', 'perpu', 'abal'	I hope this decision will not be like the fate of those who create jobs and be kicked out using the ultimate trick called fake perpu.
RT @KontraS: DANGER OF CIPTAKER PERPPU! 🚨🚨🚨 Appeal to all citizens! To raid the DPR RI Building in order to reject and revoke Perppu C ...	The dangers of the Perppu Create a call to all citizens to storm the DPR building in order to reject and revoke the Perppu C	'danger', 'perppu', 'buatkerseruan', 'to', 'whole', 'citizens', 'geruduk', 'building', 'dpr', 'ri', 'in', 'frame', 'reject', 'and', 'revoke', 'perppu', 'c'	the perppu is dangerous for all citizens to storm the DPR RI building in order to reject and revoke the perppu
RT @6an9_Said: The issuance of Perppu no. 2 of 2022 the Government's tricks are full of controversy and defiance of the Constitution https://t...	The issuance of Perppu No. 2 of 2022 is a government trick that is full of controversy and defiance of the constitution httpst	'issuance', 'perppu', 'no', 'year', 'reason', 'government', 'yang', 'laden', 'controversy', 'at once', 'disobedience', 'constitution', 'httpst'	The government's issuance of Perppu No. 2 of 2022 is full of controversy and defiance of the constitution
RT @ChaUnk_VR1: The Cipta Kerja Law has been rejected by various levels of society since it was not passed and has also resulted in a wave of layoffs ...	The job creation law has been rejected by various levels of society since it was passed and resulted in another wave of layoffs	'uu', 'creation', 'work', 'has', 'rejected', 'by', 'various', 'layers', 'society', 'since', 'not yet', 'validated', 'and', 'result', 'wave', 'layoff', 'which', 'also'	The job creation law has been rejected by various levels of society since it was not passed and resulted in a wave of layoffs
RT @DaengWahidin2: Perppu Cipta Kerja ciptakan Millions of workers have been laid off, so just lay them off, Mas Joko, without severance pay...!!! https://t.co/g23pU37...	The omnibus law on work copyright is really burdensome	'omnibus', 'law', 'create', 'work', 'it', 'burdensome', 'really'	The omnibus law on work copyright is really burdensome
RT @KontraS: To all citizens! Those who take part in the Action to Reject the Cipta Kerja Perppu today	The dangers of the Perppu Create a call to all citizens to storm the DPR building in order to reject and revoke the Perppu C	'danger', 'perppu', 'buatkerseruan', 'to', 'whole', 'citizens', 'geruduk', 'building',	the perppu is dangerous for all citizens to storm the DPR RI building in order to reject and revoke the perppu

From Table 1 is the result of the preprocessing summary on tweets that have been crawled with the keyword "Cipta Kerja" and in Indonesian. The preprocessing process uses the case folding technique which converts the text into lowercase letters, then performs the tokenizing technique, which separates each word so that it is easy to train, and uses the stopword technique which removes words that have no meaning. So that after preprocessing a word distribution is formed as shown in Figure 2 below.



Figure 2. Word distribution

From Figure 2, the distribution of words is involved which is the result of visualization of the results of preprocessing where the visualization process is formed from the frequency of words that often appear. Word distribution certainly makes it easier to identify meaningful words in determining topics. After the distribution of words is formed, then carry out sentiment distribution to make it easier to map on big data detection. The results of the sentiment distribution are shown in Figure 3.

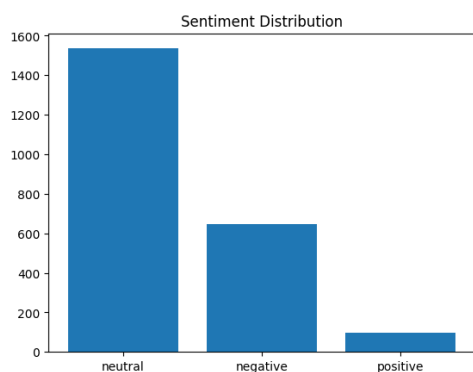


Figure 3. Sentiment distribution

Figure 3 shows that the distribution of sentiment makes it easier to form words with positive, negative and neutral meanings. These results show 150 positive sentiments, 700 negative sentiments, and 1550 neutral sentiments. After that, proceed to the detection process stage for training and testing using SVM. From the data that has been crawled and the preprocessing has been carried out, then the distribution of training and testing data is carried out. For training data it is proposed to be 70% and for testing data it is proposed to be 30% of the total data. The resulting process is visualized based on the confusion matrix as shown in Figure 4.

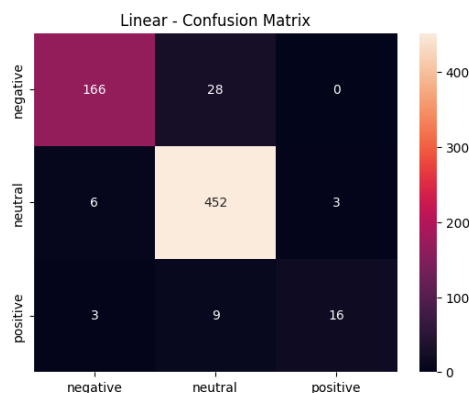


Figure 4. Linear confusion matrix

From Figure 4 it can be seen that the results of the visualization of the linear confusion matrix with the distribution of detection data are True Positive, True Negative, True Neutral and False Positive, False Negative and False Neutral. The manual calculation model for the confusion matrix in obtaining accuracy, recall, precision and F1 score can be seen in Table 2.

Table 2. Manual confusion matrix calculation table

Actual	Positive Predictions	Negative Predictions	Neutral Prediction
Positive	16	3	9
Negative	0	166	28
Neutral	3	6	452

Based on Table 2, the calculation of accuracy, precision, recall and F1 score will be carried out as follows:

- a) Accuracy

$$\text{Accuracy} = \frac{(16+166+425)}{(16+3+9+0+166+28+3+6+425)} \times 100\% = 92.7\%$$
- b) Precision

$$\text{Positive Precision} = \frac{16}{(16+3)} = 0.842$$

$$\text{Negative Precision} = \frac{166}{(166+6)} = 0.965$$

$$\text{Neutral Precision} = \frac{452}{(452+28)} = 0.924$$
- c) Recall

$$\text{Positive Recalls} = \frac{16}{(16+9)} = 0.640$$

$$\text{Negative Recalls} = \frac{166}{(166+28)} = 0.855$$

$$\text{Neutral Recalls} = \frac{16}{(452+3+6)} = 0.984$$
- d) F1-Score

$$\text{Positive F1-Score} = 2 \times \frac{(0.842 * 0.640)}{(0.842 + 0.640)} = 0.727$$

$$\text{Negative F1-Score} = 2 \times \frac{(0.965 * 0.855)}{(0.965 + 0.855)} = 0.907$$

$$\text{Neutral F1-Score} = 2 \times \frac{(0.924 * 0.984)}{(0.924 + 0.984)} = 0.953$$

After achieving accurate results, a visualization with a Receiver Operating Characteristic (ROC) curve was formed to see the successful performance of the SVM algorithm in cyberbullying cases. The ROC curve is shown in Figure 5.

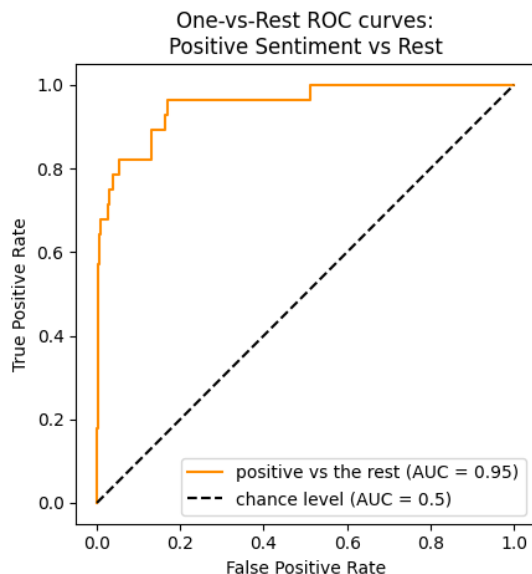


Figure 5. Results of the ROC Curve SVM model with a linear kernel

Figure 5 shows the ROC curve which is the result of the sentiment analysis performance model of tweet data with the keyword "Cipta Kerja" in Indonesian, the results of the ROC curve will produce a comparison graph between positive and negative and neutral classes in detecting big data in cyberbullying cases with SVM.

4. CONCLUSIONS

Cyberbullying detection has been successfully carried out using computational techniques on the SVM algorithm. Where the algorithm has been skilled in classifying and detecting. Of the 2,400 datasets that were crawled on Twitter social media with the keyword "Cipta Kerja" in Indonesian. Where the training and testing process achieves an accuracy of 92.7% so that further research is proposed by modifying SVM against other kernels in order to obtain optimal accuracy values. It is also possible if analyzed in a non-advanced regression technique.

ACKNOWLEDGMENT

Thanksful to the Rector of Universitas Muhammadiyah Sumatera Utara, Prof. Dr. Agussani, M.AP. and the LP2M UMSU with contract no. 100/II.3-AU/UMSU-LP2M/C/2023.

REFERENCES

[1] Chandrasekaran, S., Singh Pundir, A.K., Lingaiah, T.B. (2022). Deep learning approaches for cyberbullying detection and classification on social media. *Computational Intelligence and Neuroscience*, 2022: 2163458. <https://doi.org/10.1155/2022/2163458>

[2] Bai, Z.S., Malempati, S. (2023). An ensemble approach for cyber bullying: Text messages and images *Revue d'Intelligence Artificielle*, 37(1): 179-184. <https://doi.org/10.18280/ria.370122>

[3] Andriansyah, M., Akbar, A., Ahwan, A., Gilani, N.A.,

Nugraha, A.R., Sari, R.N., Senjaya, R. (2017). Cyberbullying comment classification on Indonesian Selebgram using support vector machine method. 2017 Second International Conference on Informatics and Computing (ICIC), pp. 1-5. <https://doi.org/10.1109/IAC.2017.8280617>

[4] Nurrahmi, H., Nurjanah, D. (2018). Indonesian twitter cyberbullying detection using text classification and user credibility. 2018 International Conference on Information and Communications Technology (ICOIACT), pp. 543-548. <https://doi.org/10.1109/ICOIACT.2018.8350758>

[5] Sainju, K.D., Young, L., Kuffour, A., Mishra, N. (2022). A machine learning and qualitative examination of cyberbullying disclosures on Twitter. *Journal of Social Media in Society*, 11(2): 209-235.

[6] Wijayanti, S.S., Utami, E., Yaqin, A. (2022). Comparison of kernels on support vector machine (SVM) methods for analysis of cyberbullying. 2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), pp. 104-108. <https://doi.org/10.1109/ICITISEE57756.2022.10057761>

[7] Kurniawan, D., Yasir, M., Venna, F.C. (2022). Optimization of sentiment analysis using naive bayes with features selection chi-square and information gain for accuracy improvement. 2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), pp. 153-160. <https://doi.org/10.23919/EECSI56542.2022.9946510>

[8] Laxmi, S.T., Rismala, R., Nurrahmi, H. (2021). Cyberbullying detection on Indonesian twitter using doc2vec and convolutional neural network. 2021 9th International Conference on Information and Communication Technology (ICoICT), pp. 82-86. <https://doi.org/10.1109/ICoICT52021.2021.9527420>

[9] Mardiyono, A., Sholihah, W., Hakim, F. (2020). Mobile-based network monitoring system using zabbix and telegram. 2020 3rd International Conference on Computer and Informatics Engineering (IC2IE), pp. 473-477. <https://doi.org/10.1109/IC2IE50715.2020.9274582>

[10] Wibowo, A., Chen, S.C., Wiangin, U., Ma, Y., Ruangkanjanases, A. (2020). Customer behavior as an outcome of social media marketing: The role of social media marketing activity and customer experience. *Sustainability*, 13(1): 189. <https://doi.org/10.3390/su13010189>

[11] Appel, G., Grewal, L., Hadi, R., Stephen, A.T. (2020). The future of social media in marketing. *Journal of the Academy of Marketing Science*, 48(1): 79-95. <https://doi.org/10.1007/s11747-019-00695-1>

[12] Patel, R., Passi, K. (2020). Sentiment analysis on twitter data of world cup soccer tournament using machine learning. *IoT*, 1(2): 14. <https://doi.org/10.3390/iot1020014>

[13] Woldemariam, Y. (2016). Sentiment analysis in a cross-media analysis framework. 2016 IEEE international conference on big data analysis (ICBDA), pp. 1-5. <https://doi.org/10.1109/ICBDA.2016.7509790>

[14] Duncan, B., Zhang, Y. (2015). Neural networks for sentiment analysis on Twitter. 2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC), pp. 275-278. <https://doi.org/10.1109/ICCI-CC.2015.7259397>

- [15] Wijaya, T.N., Indriati, R., Muzaki, M.N. (2021). Analisis Sentimen opini publik tentang undang-undang Cipta Kerja pada Twitter. *Jambura Journal of Electrical and Electronics Engineering*, 3(2): 78-83. <https://doi.org/10.37905/jjee.v3i2.10885>
- [16] Wulandari, D.A., Saedudin, R.R., Andreswari, R. (2021). Analisis sentimen media sosial twitter terhadap reaksi masyarakat pada ruu Cipta Kerja menggunakan metode klasifikasi algoritma naive bayes. *eProceedings of Engeneering*, 8(5). <https://doi.org/10.34010/komputika.v11i2.6037>
- [17] Arora, M., Kansal, V. (2019). Character level embedding with deep convolutional neural network for text normalization of unstructured data for Twitter sentiment analysis. *Social Network Analysis and Mining*, 9(1): 1-14. <https://doi.org/10.1007/s13278-019-0557-y>
- [18] Vijayarani, S., Ilamathi, M.J., Nithya, M. (2015). Preprocessing techniques for text mining-an overview. *International Journal of Computer Science & Communication Networks*, 5(1): 7-16.
- [19] Lubis, A.R., Prayudani, S., Lubis, M., Nugroho, O. (2022). Sentiment analysis on online learning during the COVID-19 pandemic based on opinions on Twitter using KNN method. *2022 1st International Conference on Information System & Information Technology (ICISIT)*, pp. 106-111. <https://doi.org/10.1109/ICISIT54091.2022.9872926>
- [20] Lubis, A.R., Prayudani, S., Fatmi, Y., Nugroho, O. (2022). Classifying news based on indonesian news using LightGBM. *2022 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, pp. 162-166. <https://doi.org/10.1109/CENIM56801.2022.10037401>
- [21] Veeranjanyulu, N., Bodapati, J.D., Buradagunta, S. (2020). Classifying limited resource data using semi-supervised SVM. *Ingénierie des Systèmes d'Information*, 25(3): 391-395. <https://doi.org/10.18280/isi.250315>
- [22] Dasari, K.B., Devarakonda, N. (2022). Detection of TCP-based DDoS attacks with SVM classification with different kernel functions using common uncorrelated feature subsets. *International Journal of Safety and Security Engineering*, 12(2): 239-249. <https://doi.org/10.18280/ijss.120213>
- [23] Masaling, N.A.P., Lubis, T., Amalia, A., Lubis, A.R., Lydia, M.S. (2022). Category based sentiment analysis for basic skin-cares using LDA and SVM approach. *2022 6th International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM)*, pp. 182-189. <https://doi.org/10.1109/ELTICOM57747.2022.10037876>
- [24] Borg, A., Boldt, M. (2020). Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Systems with Applications*, 162: 113746. <https://doi.org/10.1016/j.eswa.2020.113746>
- [25] Hamdi, M. (2022). Affirmative ant colony optimization based support vector machine for sentiment classification. *Electronics*, 11(7): 1051. <https://doi.org/10.3390/electronics11071051>
- [26] Ahsaan, S.U., Kaur, H., Mourya, A.K., Naaz, S. (2022). A hybrid support vector machine algorithm for big data heterogeneity using machine learning. *Symmetry (Basel)*, 14(11): 2344. <https://doi.org/10.3390/sym14112344>
- [27] Muhathir, Al-Khowarizmi. (2020). Measuring the accuracy of SVM with varying kernel function for classification of indonesian wayang on images. *2020 International Conference on Decision Aid Sciences and Application (DASA)*, pp. 1190-1196. <https://doi.org/10.1109/DASA51403.2020.9317197>
- [28] Kurani, A., Doshi, P., Vakharia, A., Shah, M. (2023). A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. *Annals of Data Science*, 10(1): 183-208. <https://doi.org/10.1007/s40745-021-00344-x>
- [29] Rahat, A.M., Kahir, A., Masum, A.K.M. (2019). Comparison of naive bayes and SVM algorithm based on sentiment analysis using review dataset. *2019 8th International Conference System Modeling and Advancement in Research Trends (SMART)*, pp. 266-270. <https://doi.org/10.1109/SMART46866.2019.9117512>
- [30] Styawati, S., Nurkholis, A., Aldino, A.A., Samsugi, S., Suryati, E., Cahyono, R.P. (2022). Sentiment analysis on online transportation reviews using Word2Vec text embedding model feature extraction and support vector machine (SVM) algorithm. *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*, pp. 163-167. <https://doi.org/10.1109/ISMODE53584.2022.9742906>
- [31] Kishore, D., Rao, C.S. (2020). A multi-class SVM based content based image retrieval system using hybrid optimization techniques. *Traitement du Signal*, 37(2): 217-226. <https://doi.org/10.18280/ts.370207>
- [32] Aggarwal, S., Bhatia, M., Madaan, R., Pandey, H.M. (2021). SVM prediction model interface for plant contaminates. *Traitement du Signal*, 38(4): 1023-1032. <https://doi.org/10.18280/ts.380412>
- [33] Tan, H. (2021). Machine learning algorithm for classification. *Journal of Physics: Conference Series*, 1994(1). <https://doi.org/10.1088/1742-6596/1994/1/012016>
- [34] Abdillah, W.F., Premana, A., Herdian Bhakti, R.M. (2021). Analysis of COVID-19 handling sentiment with support vector machine: Lexicon evaluation and feature extraction methods. *J. Ilm. Intech. Inf. Technol. J. UMUS*, 3(2): 160-170.
- [35] Elgammal, M.A., Mostafa, H., Salama, K.N., Nader Mohieldin, A. (2019). A comparison of artificial neural network (ANN) and support vector machine (SVM) classifiers for neural seizure detection. *2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS), Dallas, TX, USA*. <https://doi.org/10.1109/MWSCAS.2019.8884989>
- [36] Huang, S., Nianguang, C.A.I., Penzuti Pacheco, P., Narandes, S., Wang, Y., Wayne, X.U. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics and Proteomics*, 15(1): 41-51. <https://doi.org/10.21873/cgp.20063>