JURITEK: Jurnal Ilmiah Teknik Mesin, Elektro dan Komputer

E-ISSN: 2809-0799
P-ISSN: 2809-0802

Enhancing Multi-Label Hate Speech and Abusive Language Detection on Indonesian Twitter Using Recurrent Neural Networks with Hyperparameter Tuning

Tri Pratiwi Handayani[1*], Hilmansyah Gani[2]
Sistem Informasi, Universitas Muhammadiyah Gorontalo[1,2]

Alamat: Jl.Prof.Dr.Mansoer Pateda, Desa Pentadio Timur, Kabupaten Gorontalo
Tri Pratiwi Handayani : tripratiwi@umgo.ac.id

Abstract*. This study investigates enhancing multi-label hate speech and abusive language detection on Indonesian Twitter using Recurrent Neural Networks (RNNs) with hyperparameter tuning. A dataset of Indonesian tweets labeled for various hate speech and abusive language categories was preprocessed through text cleaning, tokenization, and sequence padding. A baseline RNN model was initially constructed and evaluated. Hyperparameter tuning was then performed using Keras Tuner to optimize performance. The best hyperparameters identified were an embedding dimension of 32, 32 LSTM units, and a dropout rate of 0.2. The tuned model was trained and compared with the baseline. Results indicated improved precision for labels like Abusive, HS_Group, HS_Moderate, and HS_Strong, but a decline in recall and F1-scores for labels like HS_Religion and HS_Race. Overall performance metrics showed a slight decline, highlighting trade-offs in the tuning process. In conclusion, while hyperparameter tuning can enhance certain performance aspects, it also introduces complexities and trade-offs. It is recommended to use hyperparameter tuning in model optimization with careful consideration of application requirements. Further research will explore different model architectures and additional tuning strategies for better overall performance.*

Keywords*:* Hate Speech Detection; Abusive Language Detection; Multi-label Classification; Recurrent Neural Networks (RNN); Hyperparameter Tuning

Abstrak.
Studi ini menyelidiki peningkatan deteksi ujaran kebencian multi-label dan bahasa kasar di Twitter Indonesia menggunakan Recurrent Neural Networks (RNNs) dengan penyetelan hyperparameter. Dataset tweet Indonesia yang diberi label untuk berbagai kategori ujaran kebencian dan bahasa kasar diproses melalui pembersihan teks, tokenisasi, dan penyekuan. Model RNN dasar awalnya dibangun dan dievaluasi. Penyetelan hyperparameter kemudian dilakukan menggunakan Keras Tuner untuk mengoptimalkan kinerja. Hyperparameter terbaik yang diidentifikasi adalah dimensi embedding 32, 32 unit LSTM, dan tingkat dropout 0,2. Model yang disetel dilatih dan dibandingkan dengan model dasar. Hasil menunjukkan peningkatan presisi untuk label seperti Abusive, HS_Group, HS_Moderate, dan HS_Strong, tetapi penurunan recall dan skor F1 untuk label seperti HS_Religion dan HS_Race. Secara keseluruhan, metrik kinerja menunjukkan sedikit penurunan, menyoroti trade-off dalam proses penyetelan. Kesimpulannya, meskipun penyetelan hyperparameter dapat meningkatkan aspek kinerja

tertentu, hal ini juga memperkenalkan kompleksitas dan trade-off. Disarankan untuk menggunakan penyetelan hyperparameter dalam optimalisasi model dengan pertimbangan yang cermat terhadap persyaratan aplikasi. Penelitian lebih lanjut akan mengeksplorasi arsitektur model yang berbeda dan strategi penyetelan tambahan untuk mencapai kinerja keseluruhan yang lebih baik

Kata kunci: Ujaran Kebencian; Deteksi Perundungan Siber; Klasifikasi Multi Label; Recurrent Neural Networks (RNN); Hyperparameter Tuning

INTRODUCTION

Hate speech and abusive language on social media platforms have become pressing issues, particularly in regions with high social media usage such as Indonesia. Automatic detection systems are essential to mitigate these problems. This study aims to improve multi-label classification of hate speech and abusive language on Indonesian Twitter using RNNs and hyperparameter tuning. Previous research has shown the effectiveness of various machine learning approaches, including Support Vector Machine (SVM) (Akinyemi et al., 2023; Asti et al., 2021; Aulia & Budi, 2019; Rohmawati et al., 2018), Naive Bayes (NB)(Adikara et al., 2020; Alfina et al., 2017; Putri et al., 2021a, 2021a; Rohmawati et al., 2018), and Random Forest Decision Tree (RFDT) (Hendrawan et al., 2020; Marpaung et al., 2021; Putri et al., 2021b), for detecting hate speech and abusive language on Indonesian Twitter (Ibrohim & Budi, 2019). Another study demonstrated that the Artificial Neural Network (ANN) classifier with BoW and Chi-square feature selection achieved high accuracy in multi-label classification of hate speech (Riadi et al., 2023)(Hidayatullah et al., 2023; Muhariya et al., 2023; Rawat et al., 2024). Additionally, using deep learning methods such as CNN and LSTM has been found to be effective for text classification tasks in this domain (Abidin et al., 2023; Anindyati et al., 2019; Gultom et al., 2021; Laxmi et al., 2021). This study builds upon these findings by exploring the use of Recurrent Neural Networks (RNNs)(Malik et al., 2023; Riyadi et al., 2023, 2023) and optimizing their performance through hyperparameter tuning to enhance the detection and classification of hate speech and abusive language on Indonesian Twitter.

THEORITICAL REVIEW

Several studies have focused on improving the classification of hate speech and abusive language on Indonesian Twitter using various machine learning approaches. A study by Ihsan et al. (2021) developed a method to classify tweets into abusive and hate

speech classes using a Decision Tree algorithm, with feature engineering and parameter tuning enhancing classification accuracy. Hana et al., 2020 compared SVM, deep learning, CNN, and DistilBERT for classifying hate speech, finding that SVM with Classifier Chains achieved the highest accuracy. Hendrawan et al., 2020 investigated RFDT, BiLSTM, and BiLSTM with pre-trained BERT for multi-label classification, showing the impact of different preprocessing stages. Prabowo & Azizah, 2020 used a hierarchical approach with SVM and RFDT to classify hate speech, improving accuracy over flat approaches. Lastly, Gultom et al. (2021) compared CNN and LSTM methods for classifying abusive language in Indonesian tweets, finding that CNN performed slightly better

Despite significant advancements in the detection of hate speech and abusive language on social media, several critical gaps remain in the existing literature. Firstly, while numerous studies have applied algorithms such as Support Vector Machines (SVM), Naive Bayes (NB), Random Forest Decision Trees (RFDT), Convolutional Neural Networks (CNN), and BERT models, the application of Recurrent Neural Networks (RNNs) specifically for this task has not been thoroughly explored. RNNs are particularly well-suited for processing sequential data and capturing the temporal dynamics of text, which are crucial for understanding the context and evolution of conversations on platforms like Twitter.

Additionally, the process of hyperparameter tuning, which involves optimizing the performance of machine learning models by fine-tuning parameters such as learning rate, batch size, and the number of layers, has not been given adequate attention in the context of multi-label classification of hate speech and abusive language. This study seeks to address this gap by rigorously applying hyperparameter tuning to enhance model performance, thereby achieving higher classification accuracy.

Lastly, while there is some research focused on Indonesian social media, the combination of using RNNs with hyperparameter tuning for multi-label classification in this context is new. By addressing these gaps, this study aims to contribute a more robust and accurate methodological framework for detecting and classifying hate speech and abusive language on Indonesian Twitter, thereby advancing the state of research in this critical area..

RESEARCH METHODS

The methodology of this research involves several steps: data collection, data preprocessing, model building, hyperparameter tuning, training, evaluation, and comparison.

1. Data Collection and Data Preprocessing

The data used in this study is obtained from Ibrohim & Budi (2021) and is loaded into Pandas DataFrames with ISO-8859-1 encoding. Due to limiataion of computer resources, only half data (5000) used in this study. The dataset comprises three main files:

- data.csv: Contains the raw tweets and their corresponding labels.

- abusive.csv: Contains additional abusive language data.

- new_kamusalay.csv: Contains slang words and their standard forms to aid in text normalization.

The tweets are then cleaned using a custom function that removes mentions (@usernames), URLs, non-alphabet characters, converts the text to lowercase, and strips whitespace. The Labels for multi-label classification, such as HS (hate speech), Abusive, HS_Individual, HS_Group, and others, are extracted from the dataset. The cleaned dataset is split into training and testing sets using an 80-20 split ratio. The text data is then tokenized and converted into sequences, which are padded to a maximum length of 100 tokens to ensure a uniform input size for the RNN. The experiment takes place in Google Colab using a GPU instance, which significantly accelerates the training and evaluation of the RNN models.

2. Hyperparameter Tuning

Hyperparameter tuning was performed using Keras Tuner, a library that helps automate the process of hyperparameter optimization for machine learning models. The hyperparameters tuned in this study include the embedding dimension, the number of LSTM units, and the dropout rate, which are crucial for improving the performance and generalizability of the Recurrent Neural Network (RNN) model.

Embedding Dimension (embedding_dim) is the embedding dimension defines the size of the dense vector representations of the input tokens. A larger embedding dimension can

capture more information about the relationships between words but also increases the computational complexity and risk of overfitting. In this study, the embedding dimension was tuned and an optimal value of 32 was selected, balancing the need for a rich representation of the text with computational efficiency.

Number of LSTM Units (units) is Long Short-Term Memory (LSTM) units are a type of RNN cell that can capture long-term dependencies in sequential data. The number of units determines the capacity of the network to learn from the data. Tuning the number of LSTM units is essential to ensure the model is neither too simple (underfitting) nor too complex (overfitting). An optimal value of 32 units was found, which provided a good trade-off between model complexity and performance.

Dropout Rate (dropout) is a regularization technique used to prevent overfitting by randomly setting a fraction of the input units to zero at each update during training. The dropout rate determines this fraction. A higher dropout rate can lead to a more robust model by preventing co-adaptation of neurons, but if too high, it can hinder the model's ability to learn. In this study, a dropout rate of 0.2 was selected, providing sufficient regularization to improve generalizability without significantly impacting the model's ability to learn from the data.

The best model configuration was identified through a Random Search strategy, a method where a wide range of hyperparameter values are sampled randomly and evaluated (James Bergstra & Yoshua Bengio, 2012). This approach is efficient for exploring the hyperparameter space and finding an optimal set of values. The selected configuration of embedding_dim: 32, units: 32, and dropout: 0.2 was found to yield the best performance in terms of accuracy and generalizability on the validation set. By tuning these hyperparameters, the study ensures that the RNN model is well-optimized for the task of multi-label classification of hate speech and abusive language on Indonesian Twitter.

The experiment takes place in Google Colab using a GPU instance, which significantly accelerates the training and evaluation of the RNN models. The use of GPU helps handle the computational demands of hyperparameter tuning and training deep learning models efficiently, allowing for more iterations and faster convergence to the optimal model configuration.

RESULTS AND DISCUSSION

The performance of the baseline and best models was evaluated using precision, recall, and F1-score metrics. The results are summarized in the Table 1 and Table 2.

1. Performance of Base line Model

Table 1. Evaluation of Baseline Model Classification

Label	Precision	Recall	F1-Score	Support
HS	0.81	0.76	0.78	1118
Abusive	0.84	0.86	0.85	988
HS_Individual	0.68	0.60	0.64	718
HS_Group	0.58	0.55	0.56	400
HS_Religion	0.63	0.48	0.54	164
HS_Race	0.71	0.57	0.63	118
HS_Physical	0.60	0.11	0.19	53
HS_Gender	0.67	0.04	0.07	54
HS_Other	0.75	0.67	0.71	761
HS_Weak	0.65	0.58	0.61	670
HS_Moderate	0.52	0.48	0.50	352
HS_Strong	0.73	0.64	0.68	96
micro avg	0.73	0.66	0.69	5492
macro avg	0.68	0.53	0.56	5492
weighted avg	0.72	0.66	0.69	5492
samples avg	0.42	0.39	0.38	5492
micro avg	0.73	0.66	0.69	5492
macro avg	0.68	0.53	0.56	5492

The baseline model classification report provides detailed performance metrics for each label, as well as overall metrics, based on a subset of 5,000 tweets from the larger dataset. Precision measures the proportion of true positive predictions among all positive predictions, and the baseline model shows high precision for most labels, particularly for "HS" (0.81), "Abusive" (0.84), and "HS_Other" (0.75). Recall, which measures the proportion of true positive predictions among all actual positives, is relatively high for "Abusive" (0.86) and "HS" (0.76), indicating the model's ability to identify most true positives in these categories. However, recall is notably lower for "HS_Physical" (0.11) and "HS_Gender" (0.04), suggesting the model struggles to detect these categories. The F1-scores, which balance precision and recall, are generally high for "Abusive" (0.85) and "HS" (0.78), but much lower for categories like "HS_Physical" (0.19) and "HS_Gender" (0.07). Support, referring to the number of actual occurrences of each label

in the dataset, shows that categories like "HS" (1118) and "Abusive" (988) are more prevalent, whereas "HS_Physical" (53) and "HS_Gender" (54) have low support.

Overall, the micro average, which aggregates the contributions of all labels to compute the average metric, shows precision, recall, and F1-score of 0.73, 0.66, and 0.69, respectively, indicating the model's overall performance across all labels. The macro average, which treats all labels equally, computes precision (0.68), recall (0.53), and F1-score (0.56), reflecting the model's varied performance across different labels, particularly those with lower support. The weighted average, accounting for the support of each label, shows precision (0.72), recall (0.66), and F1-score (0.69), providing a balanced view of the model's performance, giving more weight to labels with higher support. The samples average, considering the average of individual instance evaluations, has precision, recall, and F1-score of 0.42, 0.39, and 0.38, respectively, indicating the model's performance on a per-instance basis. These findings suggest that while the model is effective at detecting common types of hate speech and abusive language, it may require further optimization to improve detection of less frequent categories.

2. Best Model Classification Report

The best model classification report (Table 2) provides detailed performance metrics for each label after hyperparameter tuning, using a subset of 5,000 tweets from the larger dataset. Precision, which measures the proportion of true positive predictions among all positive predictions, improved for some categories in the best model, notably "Abusive" (0.88 from 0.84), "HS_Group" (0.61 from 0.58), "HS_Moderate" (0.56 from 0.52), and "HS_Strong" (0.75 from 0.73). However, precision for "HS_Physical" and "HS_Gender" dropped to 0, indicating the model failed to identify any true positives in these categories. Recall, which measures the proportion of true positive predictions among all actual positives, decreased for several categories, including "HS" (0.74 from 0.76), "Abusive" (0.75 from 0.86), "HS_Religion" (0.29 from 0.48), and "HS_Race" (0.30 from 0.57). These decreases suggest the model missed more true instances in these categories compared to the baseline.

Table 2. Evaluation of Best Model Classification

Label	Precision	Recall	F1-Score	Support
HS	0.81	0.76	0.78	1118
Abusive	0.84	0.86	0.85	988

HS_Individual	0.68	0.60	0.64	718
HS_Group	0.58	0.55	0.56	400
HS_Religion	0.63	0.48	0.54	164
HS_Race	0.71	0.57	0.63	118
HS_Physical	0.60	0.11	0.19	53
HS_Gender	0.67	0.04	0.07	54
HS_Other	0.75	0.67	0.71	761
HS_Weak	0.65	0.58	0.61	670
HS_Moderate	0.52	0.48	0.50	352
HS_Strong	0.73	0.64	0.68	96
micro avg	0.73	0.66	0.69	5492
macro avg	0.68	0.53	0.56	5492
weighted avg	0.72	0.66	0.69	5492
samples avg	0.42	0.39	0.38	5492
micro avg	0.73	0.66	0.69	5492
macro avg	0.68	0.53	0.56	5492

The F1-score, balancing precision and recall, showed varied results: while it remained high for "Abusive" (0.81), it dropped to 0 for "HS_Physical" and "HS_Gender," indicating a complete failure to predict these classes. Overall metrics reflected these changes: the micro average precision stayed at 0.73, but recall and F1-score decreased to 0.62 and 0.67, respectively, indicating a slight decline in overall performance across all labels.

The macro average precision (0.58), recall (0.44), and F1-score (0.49) also decreased, reflecting varied performance across different labels. The weighted average showed precision (0.71), recall (0.62), and F1-score (0.66), providing a balanced view of the model's performance with more weight given to labels with higher support. The samples average, considering individual instance evaluations, showed precision (0.38), recall (0.36), and F1-score (0.35), indicating the model's performance on a per-instance basis. These findings suggest that while hyperparameter tuning improved precision for some categories, it also introduced trade-offs, reducing recall and F1-scores for others, highlighting the need for further optimization or alternative strategies to achieve balanced performance across all categories.

SUMMARY AND RECOMMENDATION

1. SUMMARY

This study demonstrated that hyperparameter tuning can have both positive and negative impacts on model performance for multi-label classification tasks. Hyperparameter tuning was shown to enhance precision for certain labels, such as "Abusive," "HS_Group," "HS_Moderate," and "HS_Strong," indicating that the model became better at correctly identifying positive instances for these categories. However, this improvement came at a cost, as recall and F1-scores for other labels, particularly "HS_Religion," "HS_Race," "HS_Physical," and "HS_Gender," declined. The decline in recall suggests that the tuned model missed more true instances for these categories, making it less effective at identifying all relevant cases.

These results highlight the complexity and trade-offs involved in optimizing deep learning models. While tuning can improve specific aspects of model performance, it may simultaneously degrade performance in other areas. This underscores the need for a careful and balanced approach when optimizing models, taking into account the specific requirements of the application. For example, if the goal is to minimize false positives, improving precision might be desirable. Conversely, if the goal is to ensure all instances of hate speech are identified, maintaining high recall is critical.

In conclusion, while hyperparameter tuning can be a valuable tool for improving certain model metrics, it is essential to consider the broader impact on overall performance. Future work should focus on exploring additional strategies, such as combining models, employing ensemble methods, or further refining the tuning process to achieve a more balanced and robust performance across all labels. This study provides a foundation for understanding the benefits and limitations of hyperparameter tuning in the context of multi-label classification for hate speech and abusive language detection on Indonesian Twitter.

2. RECOMMENDATION

Based on the findings of this study, hyperparameter tuning is recommended as a part of the model optimization process, but with caution. It is crucial to consider the specific requirements and constraints of the application:

1. If the application demands higher precision for certain categories (for example minimizing false positives for Abusive language), hyperparameter tuning can be beneficial as it was shown to improve precision for some labels.

2. If recall is more critical (for example to ensuring all instances of hate speech are identified), further investigation is needed. The tuned model showed a decline in recall for several categories, indicating that the trade-offs might not be suitable for recall-sensitive applications.

REFERENCES

Abidin, Z., Adina, N., Arifudin, R., Purwinarko, A., Hamdani, H., & Wibisono, D. L. (2023). A deep learning approach for sentiment analysis of hate tweets. *AIP Conference Proceedings, 2614*. https://doi.org/10.1063/5.0125765

Adikara, P. P., Adinugroho, S., & Insani, S. (2020). Detection of cyber harassment (cyberbullying) on Instagram using Naïve Bayes classifier with bag of words and lexicon based features. *Proceedings of the 5th …*. https://doi.org/10.1145/3427423.3427436

Akinyemi, J. D., Ibitoye, A. O. J., Oyewale, C. T., & ... (2023). Cyberbullying Detection and Classification in Social Media Texts Using Machine Learning Techniques. *... on Computer Science …*. https://doi.org/10.1007/978-3-031-36118-0_40

Alfina, I., Mulia, R., Fanany, M. I., & Ekanata, Y. (2017). Hate speech detection in the Indonesian language: A dataset and preliminary study. *2017 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2017*, *2018-Janua*, 233–237. https://doi.org/10.1109/ICACSIS.2017.8355039

Anindyati, L., Purwarianti, A., & Nursanti, A. (2019). Optimizing Deep Learning for Detection Cyberbullying Text in Indonesian Language. *Proceedings - 2019 International Conference on Advanced Informatics: Concepts, Theory, and Applications, ICAICTA 2019*. https://doi.org/10.1109/ICAICTA.2019.8904108

Asti, A. D., Budi, I., & Ibrohim, M. O. (2021). Multi-label Classification for Hate Speech and Abusive Language in Indonesian-Local Languages. *2021 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2021*. https://doi.org/10.1109/ICACSIS53237.2021.9631316

Aulia, N., & Budi, I. (2019). Hate speech detection on Indonesian long text documents using machine learning approach. *ACM International Conference Proceeding Series*, 164–169. https://doi.org/10.1145/3330482.3330491

Gultom, R. Y., Zulkarnaen, F. I., Nurhasanah, Y., & Sholahuddin, A. (2021). Indonesian Abusive Tweet Classification based on Convolutional Neural Network and Long Short Term Memory Method. *2021 International Conference on Artificial Intelligence and Big Data Analytics, ICAIBDA 2021*. https://doi.org/10.1109/ICAIBDA53487.2021.9689728

Hana, K. M., Adiwijaya, Al Faraby, S., & Bramantoro, A. (2020). Multi-label Classification of Indonesian Hate Speech on Twitter Using Support Vector Machines. *2020 International Conference on Data Science and Its Applications, ICoDSA 2020*. https://doi.org/10.1109/ICoDSA50139.2020.9212992

Hendrawan, R., Adiwijaya, & Al Faraby, S. (2020). Multilabel Classification of Hate Speech and Abusive Words on Indonesian Twitter Social Media. *2020 International Conference on Data Science and Its Applications, ICoDSA 2020*. https://doi.org/10.1109/ICoDSA50139.2020.9212962

Hidayatullah, A. F., Kalinaki, K., Aslam, M. M., & ... (2023). Fine-Tuning BERT-Based Models for Negative Content Identification on Indonesian Tweets. *2023 8th …*. https://ieeexplore.ieee.org/abstract/document/10427046/

James Bergstra, & Yoshua Bengio. (2012). Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research*, *13*.

Laxmi, S. T., Rismala, R., & ... (2021). Cyberbullying detection on Indonesian twitter using doc2vec and convolutional neural network. *2021 9th International …*. https://ieeexplore.ieee.org/abstract/document/9527420/

Malik, V., Mittal, R., Singh, V., Mittal, A., Singh, S. V., & Diwvedi, S. P. (2023). Detection of Cyberbullying Using Modified Dense Framework. *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering, UPCON 2023*, 1181–1186. https://doi.org/10.1109/UPCON59197.2023.10434862

Marpaung, A., Rismala, R., & Nurrahmi, H. (2021). Hate Speech Detection in Indonesian Twitter Texts using Bidirectional Gated Recurrent Unit. *KST 2021 - 2021 13th International Conference Knowledge and Smart Technology*, 186–190. https://doi.org/10.1109/KST51265.2021.9415760

Muhariya, A., Riadi, I., Prayudi, Y., & ... (2023). Utilizing K-means Clustering for the Detection of Cyberbullying Within Instagram Comments. *… Des Systèmes d' …*. https://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=16331311&AN=171938951&h=jAR8LFD6hrLxqT5pNnTroYYiMkgwU7GV9uo%2FcpMzrDkpFvWMbDFoaiuRfFKcJG1xL%2BHRfUycy2tq0OOGAZ6WEA%3D%3D&crl=c

Putri, S. D. A., Ibrohim, M. O., & Budi, I. (2021a). Abusive Language and Hate Speech Detection for Indonesian-Local Language in Social Media Text. In *Lecture Notes in Networks and Systems* (Vol. 251). https://doi.org/10.1007/978-3-030-79757-7_9

Putri, S. D. A., Ibrohim, M. O., & Budi, I. (2021b). Abusive Language and Hate Speech Detection for Indonesian-Local Language in Social Media Text. *Lecture Notes in Networks and Systems*, *251*. https://doi.org/10.1007/978-3-030-79757-7_9

Rawat, A., Kumar, S., & Samant, S. S. (2024). Hate speech detection in social media: Techniques, recent trends, and future challenges. *Wiley Interdisciplinary Reviews: Computational Statistics*, *16*(2). https://doi.org/10.1002/wics.1648

Riyadi, S., Andriyani, A. D., Masyhur, A. M., Damarjati, C., & Solihin, M. I. (2023). Detection of Indonesian Hate Speech on Twitter Using Hybrid CNN-RNN. *Proceeding - International Conference on Information Technology and Computing*

2023, *ICITCOM 2023*, 352–356. https://doi.org/10.1109/ICITCOM60176.2023.10442041

Rohmawati, U. A. N., Sihwi, S. W., & Cahyani, D. E. (2018). SEMAR: An interface for Indonesian hate speech detection using machine learning. *2018 International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2018*, 646–651. https://doi.org/10.1109/ISRITI.2018.8864484