EasyChair Preprint

№ 9693

Sentiment Analysis of Local Community on G20 Performance in Indonesia Using the CNN Deep Learning Algorithm

Hilda Nuraliza, Kusumah Anggraito, Juan Ryanto, Rizka Putri Wahyuni and Muharman Lubis

February 9, 2023

Sentiment Analysis of Local Community on G20 Performance in Indonesia Using the CNN Deep Learning Algorithm

1st Hilda Nuraliza
Information Systems Department
Faculty of Industrial Engineering
Bandung, Indonesia
nuralizahilda@student.telkomuniversity.ac.id

2nd Kusumah Anggraito
Information Systems Department
Faculty of Industrial Engineering
Bandung, Indonesia
kusumahanggr@student.telkomuniversity.ac.id

3rd Juan Ryanto
Information Systems Department
Faculty of Industrial Engineering
Bandung, Indonesia
ryryanto@student.telkomuniversity.ac.id

4th Rizka Putri Wahyuni
Information Systems Department
Faculty of Industrial Engineering
Bandung, Indonesia
rizkaputriwahyuniiii@student.telkomuniversity.ac.id

5th Muharman Lubis
Information Systems Department
Faculty of Industrial Engineering
Bandung, Indonesia
muharmanlubis@telkomuniversity.ac.id

***Abstract*—The G20 Summit Forum was the first time it was held in Indonesia, many local people gave their opinions on the performance given. Given the sentiment or opinion given is quite important for the government as an evaluation material for planning in the future. The method used in this study is the text mining method with deep learning algorithms. The purpose of this study is to analyze the performance sentiment of the G20 Summit by the public using the Convolutional Neural Network (CNN) algorithm to obtain the best model pattern. Testing with a total of 5 kernels and 10 epochs obtained an accuracy value of 96%. In this case it can be proven that the model's ability to classify is considered quite good.**

Keywords— Sentiment Analysis, Deep Learning, CNN

I. INTRODUCTION (*HEADING 1*)

The G20 Summit is a forum for the governments and central bankers of 20 major economies [1]. Founded in 1999 with the aim of promoting international financial stability and driving global economic growth. The G20 Summit is a High Level Decision held annually by the G20, which is a high-level forum consisting of 19 countries and the European Union that aims to promote global economic and financial policy coordination [2]. Indonesia is a member of the G20 as well as the host of the 2022 G20 Summit and has the opportunity to demonstrate its economic and political achievements to the world [3].

The success or failure of the 2022 G20 Summit in Indonesia can be influenced by the sentiments of the people of the country itself. If the Indonesian people have positive sentiments about the 2022 G20 Summit, then the G20 2022 Summit can be said to be successful. However, if the Indonesian people have negative sentiments towards the 2022 G20 Summit, the continuation of the 2022 G20 Summit can be said to be unsuccessful [4]. Given the sentiment or opinion given is quite important for the government as an evaluation material for planning in the future.

One of the efforts to map public sentiment towards an issue or event is through a sentiment analysis process [5]. Sentiment analysis is a process of evaluating the opinions and feelings of a person or group of people on a particular topic or issue [6]. In the context of sentiment analysis on the performance of the G20 in Indonesia, we will evaluate the opinions and feelings of local people about the success or failure of the G20 in solving global problems facing Indonesia.

In this paper, we present a sentiment analysis of Indonesia's performance as the host country for the 2022 G20 Summit using a deep learning convolutional neural network (CNN) algorithm. CNN is a deep learning algorithm that is very suitable for image and text classification [7]. This algorithm is widely used in image processing and can be used for sentiment analysis by processing text data in the form of local community comments or reviews about the performance of the 2022 G20 Summit in Indonesia.

We did this research with the hope of being able to analyze both local community sentiment towards the performance of the G20 in Indonesia and also to be able to analyze how well the CNN algorithm is able to classify or predict data. In terms of accuracy (accuracy), precision (precision), recall, and f1-score and being able to analyze the success and failure rates of the model in predicting the class of a data.

II. LITERATURE REVIEW

A. *Text Mining*

Data mining is a process that employs one or more computer learning techniques (machine learning) to analyze and extract knowledge (knowledge) automatically. Data mining contains the search for the desired trend or pattern in a large database to help make decisions in the future. These patterns are recognized by certain tools that can provide a useful and insightful data analysis which can then be studied more thoroughly, which may be using other decision support tools [8].

B. *Sentiment Analysis*

Pang (2008) Sentiment analysis in a sentence describes the consideration part of the assessment of a particular entity or event [11]. Sentiment analysis can also be said as opinion mining. Sentiment analysis can be used in a variety of possible domains, from consumer products, health services, financial services, to social and political events to elections. The research trend on sentiment analysis focuses on opinions that

state or imply a positive or negative sentiment. Opinions represent almost all human activities, because opinions can affect a person's behavior. Whenever we need to make a decision, we want to know what other people think. In the real world, businesses and organizations always want to see public opinion about a product or service (Liu, 2012)[12].

C. Deep Learning

Deep Learning is a learning method that utilizes a multi-layered artificial neural network. This Artificial Neural Network is made to resemble the human brain, where neurons are connected to each other to form a very complicated network of neurons. Deep Learning or deep structured learning or hierarchical learning or deep neural is a learning method that utilizes multiple non-linear transformations, deep learning can be viewed as a combination of machines learning with AI (artificial neural network) [14]. Several algorithms are included in the Deep Learning category, including:

a. Convolutional Networks
b. Restricted Boltzmann Machine (RBM)
c. Deep Belief Networks (DBN)
d. Stacked Autocoders

D. Convulational Neural Network (CNN) Algorithm

Convolutional Neural Network (CNN) is the development of Multilayer Perceptron (MLP) which is included in the feed forward (not repeating) type of neural network. Convolutional Neural Network is a neural network designed to process two-dimensional data. CNN is included in the Deep Neural Network type because of its high network depth and is widely applied to image data [15].CNN consists of neurons that have weight, bias and activation functions.

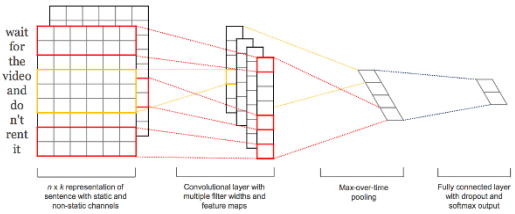

Fig 1. Neural Network model for text classification

E. Confusion Matrix

Confusion Matrix is a classifier system performance measurement built to determine the level of accuracy using performance parameters. The performance parameters used include accuracy, precision, and recall [10].

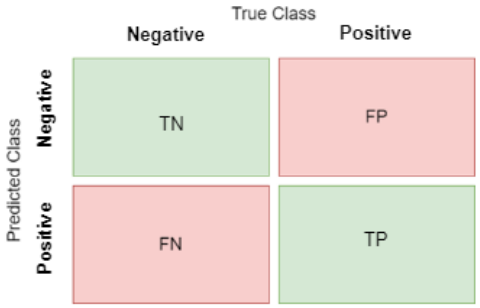

Fig 2. Confusion Matrix

Figure 1 shows that the classification results are correct according to the system decision (TP) and wrong classification according to the system (FP). Whereas documents that are not included in the category classification results, sometimes they are not members of that category (TN) and sometimes it turns out they should be members of that category (FN) [9].

1. Accuracy
   Matrix which indicates how many classes the model correctly predicted. It is the most frequently used model but in cases of unbalanced data it has a vulnerability to wrong interpretation of performance (bias). The equation of this matrix itself is as follows:

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{(II-1)}$$

2. Precision
   Matrix which shows how correctly the positive class prediction results of the model are positive (actual) class. However, this matrix cannot clearly describe how many positive (actual) predicted results have. The equation of this matrix itself is as follows:

$$Presisi = \frac{TP}{TP+FP} \quad \text{(II-2)}$$

3. Recall
   Matrixwhich indicates how correctly the (actual) positive class is predicted as positive class. However, this matrix cannot describe how well the model predicts positive classes as positive (actual) classes. The equation of this matrix itself is as follows:

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

4. F1 Score
   Matrixwhich makes up for the lack of precision and recall in assessing the performance of the positive class by calculating the harmonic average of the two. However, because the two previous matrices only focused on the positive class, the f1 score also could not describe specifically the performance assessment of the negative class. However, all of these things are overcome by implementing a weighted version that takes into account all the existing classes and their distribution. The equation of this matrix itself is as follows:

$$F1 - Score = 2 * \frac{presisi * recall}{presisi + recall} \quad \text{(II-4)}$$

III. Prepare Your Paper Before Styling

In the research used in the text mining process, the following are the steps used in doing sentiment analysis using the CNN algorithm. These steps begin with data collection, data preprocessing, data labeling, ABSA, and end with an evaluation. Figure 3 below shows the steps of text mining, namely as follows.

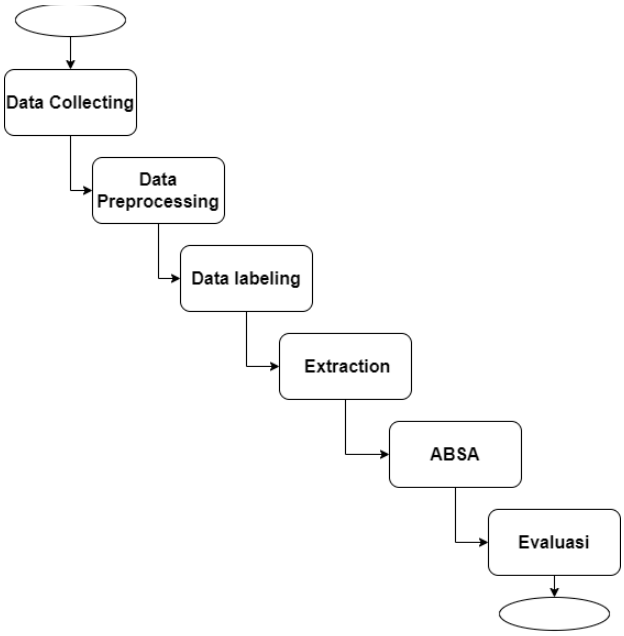

Fig 3. Research steps

A. *Data collection*

Data collection was carried out by taking reviews from the Twitter social media platform and storing the data in a file with the .csv extension. After the data is obtained, the data is separated into sentences for further processing. The review data will be used as training data for model building and data testing to see the final results of the classification and accuracy tests. The selected data is primary data, which means that the data has been collected by the first party, namely in the form of a review dataset on the performance of the G20 obtained from the Twitter social media platform. The dataset consists of 747 rows of data and 2 attribute columns after the process of merging the entire dataset from reviews with labels on sentiment indicating good or bad review results as in Table 1.

Table 1. Data Description

No	attribute	Row
1.	Tweets	747

The data is data ready to be processed and classified based on positive, neutral and negative sentiments, so that further processing can be carried out, namely the data preprocessing process.

B. *Data Preprocessing*

As the theory and research methods have been described in Figure 3. the second stage of data mining is data preprocessing, at this stage the author performs cleaning and refinement of the data to ensure whether the data to be used is of good quality. In this process, there are six things that need to be ensured, namely data accuracy, completeness, consistency, timeliness, reliability, and can be interpreted properly. In this study the authors used the Google Colab platform as a data preprocessing medium. The several stages of data preprocessing carried out by the author are shown in Figure 4.

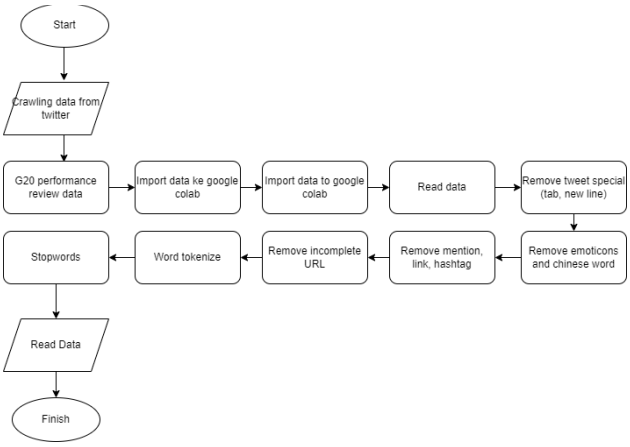

Fig 4. Data preprocessing

C. *Data Labeling*

The next stage is data labeling, labeling is done on the training data after the data has been obtained. The necessary data labeling is sentiment labeling and also category aspects. The review sentiment labeling is divided into 3, positive, neutral and negative. The data labels used to train the model are 747 data labeled automatically using the Text Blob library. Each of these data has one aspect label, and one sentiment label. The amount of data on each label is not the same, so the data is imbalanced, therefore an oversampling technique is used to make the data in each class balanced using the SMOTE method.

	tweet	label
0	buat indonesia mendunia	Positive
1	Sukses dilaksanakan di indonesia	Positive
2	Bangga dengan Pak Jokowi ampu menjadikan Indon...	Positive
3	Momen Kedatangan Para Pemimpin Di Bali	Positive
4	Presidensi kado terbaik bagi Indonesia	Positive
...	...	...
743	sukses indonesia mendunia	Positive
744	Kaleidoskop Euforia Presidensi Indonesia di T...	Positive
745	Dukung kepemimpinan Indonesia pada Presidensi	Positive
746	Kita patut berbangga krn kita berada di puncak...	Positive
747	Pulih Bersama	Positive

Fig 5. Data labeling

D. *Extraction Feature*

At this stage, feature extraction is performed using negation handling and POS tagging. Negation Handling, namely combining words which contain one of the words "not", "less", "don't", or "not" followed by the next word. POS tagging (part-of-speech tagging) is the process of marking words in a text/corpus according to certain parts of speech, based on their definition and context. With POS Tagging you can categorize word classes into nouns, adjectives, verbs, etc. [19]. Feature extraction is done for each aspect and sentiment.

E. *Aspect Based Sentiment Analysis (ABSA)*

The CNN method uses several layers, namely embedding, convolution, pooling, and a dense or fully connected (FC) layer. The embedding layer is used to provide a dense

representation of words and their relative meanings. Embedding will map each word into a continuous vector space so as to form a continuous and distributed vocabulary. Convolution layers are used for feature extraction. In it, calculations are carried out to obtain an output called a feature map, in order to obtain a feature map that represents the input characteristics. The pooling layer progressively reduces the spatial size of the representation to reduce the number of parameters and computations in the network. The FC layer acts as the output layer of the artificial neural network [20]

F. Evaluation

Evaluation is done by using hold-out method and confusion matrix. The hold-out method divides the data into 2 parts, namely training data and test data. The training data is used for Data Collection Preprocessing ABSA Feature Extraction Data Labeling Evaluation of training the model. Validation data is used to see how well the model is performing. Confusion matrix is commonly used to measure the performance of a classification algorithm [20].

IV. Results & Discussion

In this section we will discuss training the Word2Vec model to get word embedding, training the CNN model, testing the model and analyzing the results.

A. Training Word2vec

Word2vec model training is carried out to obtain word embedding that will be used in the sentiment classification process later. The training for the Word2vec model in this study was carried out using a collection of sentiment reviews related to the performance of the G20 implementation in Indonesia on Twitter in the Indonesian language (consisting of 747 sentences). Then, the word embedding that has been obtained is fine-tuned with the words in the suggestion data, so that the word embedding that is formed is more appropriate to the context. The word embedding dimension used in this study is 64 dimensions. In Table 2 is a sentiment analysis model scenario while in Fig. 6 shows the model architecture.

Table 2. Scenario Model

Model	Method	Kernels	Hidden_ Dims	Dropouts
squential	CNN	5	64	0.7

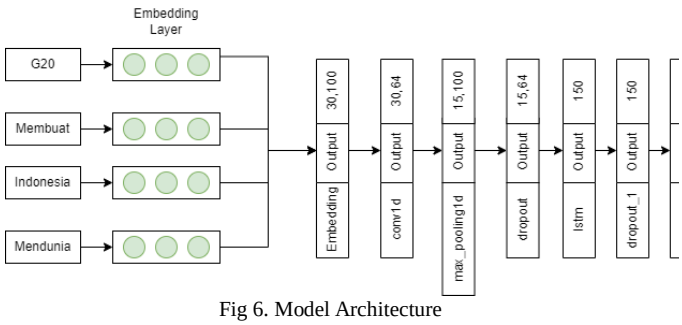

Fig 6. Model Architecture

B. Training Scheme

In CNN with this architecture, the hidden size used is 64, the activation function used is relu, the number of kernels is 5, the max-pooling size is 2, the optimizer used is adam (learning rate=0.001). CNN also implements a dropout of 0.7. The number of epochs used is 10 and the batch size is 32. All processes in this sub-chapter are implemented using the keras library.

C. Test result

Testing is carried out using the confusion matrix evaluation, namely the results of the accuracy of the resulting model predictions.

Table 3. Test Results

network	Epoch	Precision	recall	accuracy	F1-Score
CNN	10	0.97	0.98	0.96	0.97

Table 3 describes the test results obtained from the classification of the 3 sentiment classes using CNN with the kernel used, namely 5. From the test results in table 3, it can be seen that the CNN model used works well, because there is no big difference between the accuracy of the data training and data evaluation in each epoch. Accuracy increased stably from epoch 1 to epoch 2 to epoch 8 accuracy did not experience a significant increase and tended to stagnate, finally at epoch 9 and 10 the accuracy decreased. Based on some of the tests that have been carried out above, it can be concluded that CNN can work very well in conducting sentiment analysis of public review data on the performance of the G20 implementation in Indonesia. Even though the student suggestion data used in this study can contain very long sentences, reaching up to 5000 words, CNN can produce very good sentence representations, so that the performance of sentiment analysis is also better. CNN is capable of producing an accuracy of 0.96 or equivalent to 96%.

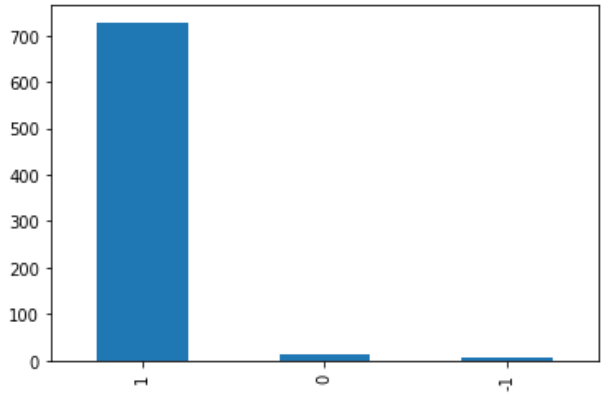

Fig.7 Label distribution graph

In Fig.7 it can be seen that the exploration results show the distribution of positive, neutral and negative class label features. It can be seen that the distribution of feature labels in the positive class has the most frequency compared to the neutral and negative sentiment classes.

Conclusion

Based on the results of the tests that have been carried out with the trial use of the CNN algorithm for the analysis of problem sentiment as follows.

1. The accuracy results obtained by the CNN algorithm model for text sentiment with kernel 5 and epoch 10 are 96%, it can be concluded that the model produced in making predictions on test data is quite good.

2. The labeling results automatically obtained positive sentiment having the highest number of frequencies. It can be concluded that based on public opinion the performance of the G20 implementation in Indonesia was very good

REFERENCES

[1] Nelson, RM (2010). G-20 and International Economic Cooperation: Background and Implications for Congress. DIANE Publishing

[2] Wouters, J., & Ramopoulos, T. (2012). The G20 and Global Economic Governance: Lessons form Multi-Level European Governance. J. Int'l Econ. L., 15, 751.

[3] SALYA, S., & NANDANA, Y. (2022). Indonesia's 2022 Presidency for the G20 Summit in a Strategic Intelligence Perspective Anticipating Security and Hidden Agenda Issues. International Journal of Environmental, Sustainability, and Social Science, 3(2), 288-298.

[4] Wicaksana, IGW (2022). How does Indonesia exercise agency in the contested and complex regional environment?. The Pacific Review, 35(2), 297-318.

[5] Nazir, A., Rao, Y., Wu, L., & Sun, L. (2020). Issues and challenges of aspect-based sentiment analysis: a comprehensive survey. IEEE Transactions on Affective Computing.

[6] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.

[7] Zhao, W., Joshi, T., Nair, VN, & Sudjianto, A. (2020). Shap values for explaining CNN-based text classification models. arXiv preprint arXiv:2008.11825.

[8] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[9] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interfaces," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[10] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[11] Mandias, FA, Application of Data Mining for Evaluation of Student Academic Performance at Klabat University Using the Classification Method, National Conference on Systems & Informatics 2015, STIKOM Bali, 9-10 October 2015, 2015

[12] Han, J., and M.Kamber, Data Mining Concepts and Techniques Second Edition, Elsevier, San Frasisco, 2006

[13] Nuraliza, H., Pratiwi, ON, & Hamami, F. (2022). Sentiment Analysis of IMBd Film Review Dataset Using Support Vector Machine (SVM) and Feature Importance Selection. Journal of Mirai Management, 7(1), 1-17.

[14] Pang, Bo., and Lee, Lillian. 2008. "Opinion Mining and Sentiment Analysis" Computer Science Department

[15] Widiarsono, Teguh 2005. "Practical Matlab Learning Tutorial". pdf, Jakarta

[16] Saputra, CB, Muzakir, A., & Udariansyah, D. (2019). Analysis of Public Sentiment Against # 2019 replacepresident Based on Opinions From Twitter Using the Naive Bayes Classifier Method. In Bina Darma Conference on Computer Science (BDCCS) (Vol. 1, No. 2, pp. 403-413).

[17] Primartha, Rifkie. Learning Machine Learning Theory and Practice. Palembang, 2018.1

[18] Suartika E. P, Arya Yudhi Wijaya Wijaya, and Rully Soelaiman . "Image Classification Using Convolutional Neural Network (Cnn) on Caltech 101." 2016

[19] Nugroho, PA, Fenriana, I., & Arijanto, R. (2020). Implementation of Deep Learning Using Convolutional Neural Network (Cnn) on Human Expression. Algor, 2(1), 12-20.

[20] Yao, C., Qu, Y., Jin, B., Guo, L., Li, C., Cui, W., & Feng, L. (2016). A convolutional neural network model for online medical guidance. IEEE Access, 4, 4094-4103.

[21] R. Siriringoringo, "Unbalanced Data Classification Using SMOTE and K-Nearest Neighbor Algorithms," J. ISD, vol. 3, no. 1, pp. 44–49, 2018

[22] Nguyen, M. (2018). Illustrated Guide to LSTM's and GRU's: A step by step explanation. Retrieved September 10, 2019, fromhttps://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-astep-by-step-explanation-44e9eb85bf21

[23] Nayoan, RAN (2019). Feature-Based Sentiment Analysis on Tourist Place Reviews Using the Convolutional Neural Network (CNN) Method (Doctoral dissertation, Islamic University of Indonesia).

[24] Zhang, M., Li, W., & Du, Q. (2018). Diverse region-based CNN for hyperspectral image classification. IEEE Transactions on Image Processing, 27(6), 2623-2634.

[25] Guo, S., Höhn, S., Xu, F., & Schommer, C. (2019). Personalized sentiment analysis and a framework with attention-based Hawkes process models. In International conference on agents and artificial intelligence (pp. 202-222). Springer, Cham.