DOI:10.14421/ijid.2021.3007 IJID (International Journal on Informatics for Development), *e-ISSN: 2549-7448*
Vol. 10, No. 2, 2021, Pp. 82-91

Sentiment Analysis of Tweets on Prakerja Card using Convolutional Neural Network and Naive Bayes

Pahlevi Wahyu Hardjita
Informatics Department
Faculty of Science and Technology
UIN Sunan Kalijaga
Yogyakarta, Indonesia
pahleviwahyu@gmail.com

Nurochman
Informatics Department
Faculty of Science and Technology
UIN Sunan Kalijaga:
Yogyakarta, Indonesia
nurochman@uin-suka.ac.id

Rahmat Hidayat
Informatics Department
Faculty of Science and Technology
UIN Sunan Kalijaga
Yogyakarta, Indonesia
rahmat.hidayat@uin-suka.ac.id
Jacobs University, Campus Ring 1, 28759
Bremen, Germany
r.hidayat@jacobs-university.de

Article History
Received Sept 2nd, 2021
Revised Dec 30th, 2021
Accepted Dec 30th, 2021
Published Jan, 2022

***Abstract*—The Indonesian government launched the Prakerja (pre-employment) card in the midst of the COVID-19 pandemic, and the local citizens have voiced their opinions about this controversial program through social media such as Twitter. People's comments on it can be useful information, and this research tries to analyze the sentiment regarding the Prakerja Card program using the Convolutional Neural Network and Naive Bayes methods. The main task in this sentiment analysis is analyzing the data and then classifying them into one of the following classes: positive, negative or neutral. Naive Bayes is an algorithm that is often used in sentiment analysis research, and the results have been very good. Convolutional neural network (CNN) is a deep learning algorithm that uses one or more layers commonly used for pattern recognition and image recognition. Having applied these methods, this research found that the CNN model with the GlobalMaxPooling layer is the best model of the other two CNN models. Sentiment analysis has the best accuracy of 78.5% on the CNN method, and NBC of 76.2% accuracy. The best accuracy result in K-fold with five classes is 85.4% on the CNN model with a learning rate optimization of 0.00158. While the average accuracy on NBC only reached 75.3%.**

Keywords*—*classification; deep learning; sentiment analysis; prakerja card; twitter

1 INTRODUCTION

In the last few months, Indonesia and even the world have been hit by an outbreak caused by the coronavirus known as COVID-19. This outbreak has been designated by the World Health Organization (WHO) as a pandemic. Due to the social restriction policy such as PSBB or recently called PPKM, approximately 2.1 million workers in Indonesia have been laid off. [1] Government intervention in the fate of workers affected by the pandemic was urgently needed because laborers could not survive without income. This pandemic has dramatically impacted various sectors such as tourism, education, agriculture, industry, and hospitality. In the midst of the crisis, the government launched the Prakerja Card program. It is a competency development program in the form of financial assistance for job seekers, laid-off workers, or workers who need competency improvement. Not all Indonesian people get this opportunity [2]. The Coordinating Ministry for Economic Affairs of the Republic of Indonesia stated that they prioritize laid-off workers or workers as well as micro and small business actors directly affected by the COVID-19 pandemic. Applicants who pass the selection are offered the training to develop competencies at the workplace and entrepreneurship. [3]

Many people react to the Prakerja Card program, and they voice their opinions through social media such as Twitter. Their comments can be valuable data which can be analyzed using sentiment analysis. A number of sentiment analysis research have been conducted since around 2002 and become an essential discussion in the field of Natural Language Processing. Sentiment analysis, also known as opinion mining, aims to study opinions, sentiments, and emotions expressed in texts. One of the sciences that deal with this is data mining. Data mining has several branches of science that explore data according to the context of the existing data, one of which is text mining. Text mining is the process of finding information in a collection of documents and automatically identifying patterns that are formed and associated with information obtained from unstructured data sets [4]. Text classification is the process of finding similarities in documents, corpus, or groups of previously labeled documents (supervised learning) based on topics and themes shown by the document collection [5].

The main objective in sentiment analysis is to analyze the data and then classify them into one of the following classes: positive, negative or neutral [6]. This study uses a supervised learning approach with the Convolutional Neural Network (CNN) algorithm, a variant of the Artificial Neural Network, included in Deep Learning, and also applies to Naïve Bayes. Convolutional neural network (CNN) is a class of artificial neural networks that have been predominantly used in computer vision tasks [7], [8]. Naïve Bayes Classifier is a classification method rooted in Bayes theorem with classification methods using probability and statistical methods [9]. Naive Bayes is an algorithm that is often used in sentiment analysis research. Convolutional neural network (CNN) is a deep learning algorithm that uses one or more layers commonly used for pattern recognition and image recognition. In this research, Convolutional neural network (CNN) is used for pattern recognition in the text or, in this case, sentiment analysis regarding the government's Prakerja Card program. Compared to Naive Bayes which is a standard machine learning method for sentiment analysis, the result from CNN is better because it has many architectures such as LSTM and Bidirectional LSTM. Long Short-Term Memory (LSTM) is a technique commonly used in natural language processing that allows the model to understand the meaning of a sentence based on word order [10]. Compared to other layers, the LSTM layer works better when combined with the CNN model for sentiment analysis.

2 METHOD

The method used in this research is the library research. This section will describe the activities associated with the method of collecting library data, reading and processing research materials.

2.1 The Object of Research

The research object in this study was the tweets about Prakerja Card by the Twitter users. The keyword in this search is "prakerja" using the API from Twitter. The data were taken when the issue was trending on Twitter with the hashtag "prakerja", between May 2020 and October 2020 with random data collection times.

2.2 The Method of Research

This study used a supervised learning approach with the Convolutional Neural Network (CNN) and Naïve Bayes algorithms. CNN is a class of artificial networks that has been predominantly used in computer vision tasks. Naïve Bayes Classifier is a classification based on the Bayes theorem with classification methods with probability and statistical methods.

2.2.1 *Convolutional Neural Network (CNN):* a machine learning method from the development of Multi-Layer Perceptron (MLP). It is designed to process two-dimensional data [11]. CNN is included in the type of Deep Neural Network because of the depth of the network level and is widely implemented in image data. CNN has two methods: the classification using feedforward and the learning stage using backpropagation.

2.2.2 *The Long Short-Term Memory (LSTM)*: architecture developed as a solution to the vanishing gradient problem encountered in conventional RNNs. Vanishing gradient happens because the gradient is getting smaller until the last layer makes the weight value unchanged, making it unable to get a better result or converge. On the other hand, the increasing gradient causes the weight values in several layers to also increase so that the optimization algorithm becomes divergent or is called an exploding gradient. LSTM consists of the following steps:

- Forget Gate: In this section, information that is less needed or has little meaning for the processed case will be removed using the sigmoid function.
- Input Gate: This process will sort and determine certain information that will be updated to the cell state section by using the tanh activation function.
- Cell State: Update the old cell state value ct-1 to the new cell state ct.
- Output Gate: By using sigmoid to generate the output value in the hidden state and placing the cell state at tanh. After generating the sigmoid output value and the tanh output value, the two activation results are multiplied before going to the next step.

2.2.3 *Naïve Bayes Classifier:* is a statistical approach to perform inductive inference on classification problems. This method uses conditional probability as the basis.[6] In statistics, conditional probability is expressed as Probability X in Y is the interaction probability X and Y from probability Y. Bayes' theorem can be written in the following equation.

2.2.4 *TF-IDF:* Term frequency focuses more on terms that often appear in a document, while Inverse Document Frequency focuses more on giving low weight to terms that appear in many documents [12]. Term frequency (TF) states the frequency of occurrence of words in each given document indicating how important the word is in each of these documents. At the same time, Document Frequency (DF) is the frequency of documents containing the frequency indicating how common the word is. IDF is the inverse of the DF value.

2.2.5 *K-fold:* is a cross-validation method used to estimate the capabilities of a machine learning model on invisible data. It is usually used to validate the model because it is easy to understand and to implement, and the results have a higher informative value than ordinary Validation Methods. How k-fold works and the training data are separated into many parts randomly, then each part will be used as training data. Data tests are carried out in each experiment.

2.3 Procedures

The data used in this study included data crawled from tweets between May 2020 and October 2020 with random data collection times. The keyword was "prakerja." In the process of collecting data using the Twitter API, the data that could be collected were 4875 tweets. The data were separated into 4725 training data and validation and 150 testing data for the automatic classification process using the model that had been made. The following is the flowchart of sentiment analysis for this research in Fig.1.

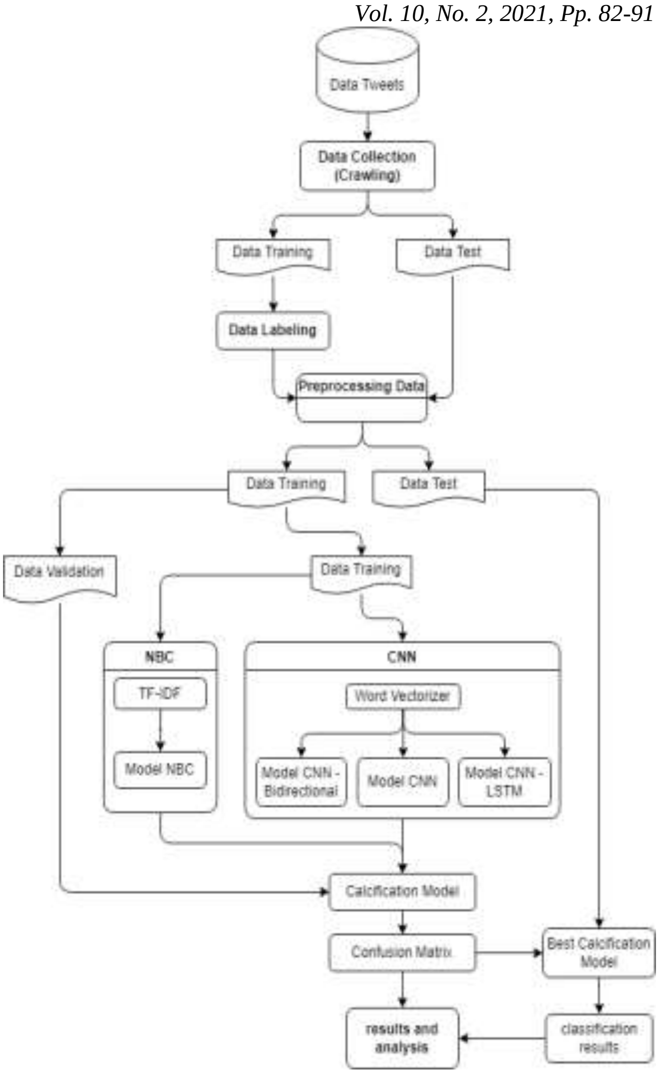

Figure 1. Flowchart of sentiment analysis

3 RESULT AND DISCUSSION

Several stages that are carried out in this research are as follows:

3.1 Data Collection

In collecting data using the Twitter API, this research found 4875 tweets. The data were obtained by crawling Twitter data using the API from Twitter with the Tweepy library from the Python programming language. The files obtained were in Excel and CSV documents containing raw data. The file contained data of 4875 tweets. There was no special processing method for the tweets. Only tweets in Indonesian language about "prakerja" were processed. These tweets could be in the form of regular tweets, news, or retweets. All forms of tweets were preprocessed so that they had similar form. It was done by removing the word "RT" in retweet text or removing links and stop words in news tweets. The examples of raw data samples can be seen in the following Table 1.

Table 1. Sample Data

No	Tweet
1	Modal Nomor KTP dan Data Diri Bantuan Pemerintah Rp 3,55 Juta Ditransfer, Kapan Kartu Prakerja Dibuka Lagi? https://t.co/dN2lebw1I8
2	RT @_BayuSapta: Sok-sok an Anti Presiden Anti Pemerintahan ,tapi Status.e "Alhamdulillah Dana Prakerja Cair" Wes Taek tok kon , Rupamu Ko…
3	@girza_peach @BPJSTKinfo Tapi setau saya tidak ada peraturan yang menjelaskan kalau ikut program prakerja gak dapat BSU BPJSTK ini

3.2 Data Labelling

Data labeling is the step of preparing training data by labeling and grouping the tweets into three classes: positive, negative, and neutral. At the data labeling stage, labeling was carried out. In this process, the author did not do it himself in order to avoid personal subjectivity. An example of the labeling results can be seen in Table 2.

Table 2. Example of Labeled Tweets

No	Tweet	Label
1	Modal Nomor KTP dan Data Diri Bantuan Pemerintah Rp 3,55 Juta Ditransfer, Kapan Kartu Prakerja Dibuka Lagi? https://t.co/dN2lebw1I8	neutral
2	RT @_BayuSapta: Sok-sok an Anti Presiden Anti Pemerintahan ,tapi Status.e "Alhamdulillah Dana Prakerja Cair" Wes Taek tok kon , Rupamu Ko…	negative
3	@girza_peach @BPJSTKinfo Tapi setau saya tidak ada peraturan yang menjelaskan kalau ikut program prakerja gak dapat BSU BPJSTK ini	positive

At this stage, this research found 2036 positive labels, 1730 neutral labels, and 959 negative labels.

3.3 Data Preprocessing

The data preprocessing stage is a stage to prepare the data that have been obtained from the data labeling stage, in order that the data be processed optimally at the next step [6]. It must be done because data mining requires consistent data. This affects the quality of the results. In the tweet data generated, a data selection process was still needed to be more consistent, which affected the next stage [12]. Therefore, it is necessary to carry out data preprocessing stages in Fig.2 as follows:

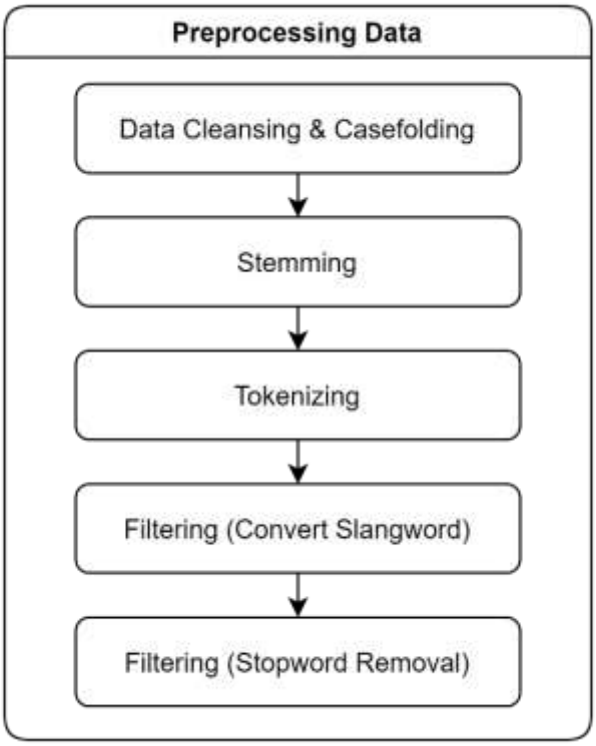

Figure 2. Flowchart of preprocessing data

3.3.1 *Case Folding and Cleansing:* This is a step to convert the entire text in the document into a standard form like lowercase. Meanwhile, cleansing is a step to clean tweet data from punctuation, URL, username, hashtag, Unicode, and also characters other than letters.

3.3.2 *Stemming*: It is a step to reduce the number of different indexes from a document and to group other words that have the same standard word and meaning but have a different form because they get different affixes.

3.3.3 *Tokenizing*: It is a step of truncating an input string based on each word that composes it or a stage of dividing the sentence text into word tokens; usually, the division is based on spaces or punctuation marks in sentences.

3.3.4 *Filtering*: It is a step to take important words from the token results. It can use the stoplist algorithm (remove less important words) or wordlist (save important words). Stoplists/stopwords are non-descriptive words that can be discarded in the bag-of-words approach.

The following is the result of the data preprocessing stage presented in Table 3.

Table 3. Data Preprocessing Result

Tweet	Result
Modal Nomor KTP dan Data Diri Bantuan Pemerintah Rp 3,55 Juta Ditransfer, Kapan Kartu Prakerja Dibuka Lagi? https://t.co/dN2lebw1I8	modal nomor ktp data bantu perintah rp juta transfer kartu prakerja buka
RT @_BayuSapta: Sok-sok an Anti Presiden Anti Pemerintahan ,tapi Status.e "Alhamdulillah Dana Prakerja Cair" Wes Taek tok kon , Rupamu Ko…	soksok an anti presiden anti perintah status alhamdulilah dana prakerja cair tai tok kon rupa ko
@girza_peach @BPJSTKinfo Tapi setau saya tidak ada peraturan yang menjelaskan kalau ikut program prakerja gak dapat BSU BPJSTK ini	tidak atur program prakerja bsu tidak bpjstk

After the preprocessing stage, the amount of data was reduced to 4718 because some tweets after the preprocessing data became empty text or null.

3.4 *Analysis*

Data that had gone through the labeling and preprocessing were processed using Naive Bayes Classifier (NBC) and Convolutional Neural Network (CNN) algorithms. The feature extraction process was carried out using TF-IDF for the NBC model and Text Vectorizer and Embedding from TensorFlow for the CNN model.

In the TF-IDF method, Term Frequency focuses more on terms that often appear in a document, while Inverse Document Frequency focuses more on giving low weight to terms that appear in many documents [13]. Term frequency (TF) states the frequency of occurrence of words in each given document indicating how important the word is in each of these documents. At the same time, Document Frequency DF is the frequency of documents containing the word indicating how common the word is. IDF is the inverse of the DF value.

The result of the TF-IDF is the product of the TF multiplied by the IDF. Word weights are greater if they frequently appear in a document and smaller if they appear in many documents. TF-IDF can be formulated as follows, which are given in Equation 1 and 2. In the calculation, the TF-IDF algorithm calculates the weight of the document [14] :

$$W_{t,d} = tf_{t,d} \times idf_t \quad (1)$$

$$W_{t,d} = tf_{t,d} \times (\log\frac{N}{df_t} + 1) \quad (2)$$

The following is an example of TF-IDF calculations presented in Table 4:

Table 4. Example of TF-IDF Calculation

Term	***TF***		***IDF***	***TF-IDF***	
	D1	**D2**		**D1**	**D2**
materi	1	0	1.69897	1.69897	0
Program	1	2	1.22185	1.22185	2.4437
kartu	0	1	1.69897	0	1.69897

The Text Vectorizer feature extraction is different from the TF-IDF feature extraction. It changes terms or sentences into word indexes and then produces a series array representing the index of each term or word. The following is an example of Word Index for Text Vectorizer presented in Table 5.

Table 5. Examples of Word Index

Index	***Term***
1	-
2	prakerja
3	program
4	kartu
5	tidak
6	laksana

After completing feature extraction, training and data testing with a data split ratio of 7:3 were conducted using the following model.

3.4.1 *Naïve Bayes Classifier (NBC)*

In this study, the Naive Bayes Classifier (NBC) used multinomial Naive Bayes from the library Skit-learn, one of NBC used in text classification with fixed parameter settings. The data were represented as a vector sum TF-IDF [15]. The distribution was parameterized by a vector for each class, the number of features in the text classification in the form of vocabulary size, and the probability of features appearing in the sample belonging to the class. In the calculation, Multinomial Naive Bayes was the Naive Bayes algorithm that manages Multinomial data in text classification. Data in Multinomial Naive Bayes was represented as a total vector, so Multinomial Naive Bayes was suitable for estimating the frequency of words in a document. In Multinomial Naive Bayes, first the probability of words in a class (prior) was calculated as the following equation.

$$P(c|d) \propto P(c) \prod_{1 \le k \le n_d} P(t_k|c) \quad (3)$$

In the Equation 3 $P(t_k|c)$ is the conditional probability of the term t_k occurring in document d of class c. P(c) is the initial probability of a document occurring in class c. The probability of documents d in c is done using the following equation.

$$\hat{P}(c) = \frac{N_c}{N} \tag{4}$$

In the Equation 4 N_c is the number of documents in class c and N is the total number of documents. The conditional probability $P(t|c)$ as the relative frequency of terms in the document belonging to the class, as in the following equation.

$$\hat{P}(t|c) = \frac{T_{tc}}{\sum_{t' \in V} T_{tc'}} \tag{5}$$

In the Equation 5 T_{tc} is the number of occurrences of vocabulary t in the training document from class c. $\sum_{t' \in V} T_{tc'}{}'$ is the sum of all terms in all documents in class c including redundant terms in the same document. The limited vocabulary in the document results in the estimated frequency of $P(t|c)$ equal to zero (0), so one or laplace smoothing is added as the following Equation 6.

$$\hat{P}(t|c) = \frac{T_{tc}+1}{\sum_{t' \in V}(T_{tc'}+1)'} = \frac{T_{tc}+1}{(\sum_{t' \in V} T_{tc'})+B'} \tag{6}$$

Where B = | V | is the number of terms in the vocabulary in the training data [16]. The following is the result of a confusion matrix for the multinomial Naive Bayes algorithm presented in Table 6.

Table 6. Confusion Matrix of the Implementation of Naive Bayes

NBC	Predict Class		
True Class	Negative	Neutral	Positive
Negative	197	36	67
Neutral	22	318	149
Positive	6	57	564

Then from the confusion matrix above, Equation 7-10 were used to calculate performance results such as Accuracy, Precision, Recall, and F1-Score from the model.

- *Accuracy*

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FN+TN)} \tag{7}$$

- *Precision*

$$Precision = \frac{TP}{(TP+FP)} \tag{8}$$

- *Recall*

$$Recall = \frac{TP}{(TP+FN)} \tag{9}$$

- *F1 Score*

$$F1\ Score = \frac{2*(Recall * Precision)}{(Recall + Precision)} \tag{10}$$

The results of calculating accuracy, precision, recall, and f1 score using the confusion matrix in this model are shown in Table 11.

3.4.2 *Convolutional Neural Network (CNN)*

For the models, Convolutional Neural Network (CNN) was made into three pieces models of CNN - Bidirectional LSTM, GlobalMaxPooling CNN, and CNN - LSTM. Our parameters are presented in Table 7.

Table 7. Parameters of CNN

Dimension of layer Embedding	In: 5000 Out: 256
Dropout	0.5
Kernel Size	8
Learning Rate	0.001
Max Epoch	20
Callback	Early Stop
Optimizer	Adam

The use of Early Stop and Dropout layers was helpful to prevent the model from being too focused on training data, so it cannot make correct predictions. Max Epoch was useful for limiting the iterations carried out during the training process. Then Early Stop stopped the training process immediately even though it had not reached the iteration limit. The CNN model layer structure is illustrated in Table 8 as follows.

Table 8. Model Structure of CNN

Bidirectional LSTM	***GlobalMaxPooling***	***LSTM***
Embedding (256)	Embedding (256)	Embedding (256)
Conv1D (128)	Conv1D (128)	Conv1D (128)
Bidirectional(LSTM)	GlobalMaxPool1D	LSTM (64)
Dense (32)	Dense (32)	Dense (32)
Dropout (0.5)	Dropout (0.5)	Dropout (0.5)
Dense (3, Softmax)	Dense (3, Softmax)	Dense (3, Softmax)

The following is the result of a confusion matrix for the all model Convolutional Neural Network (CNN) algorithm and presented in Table 9.

Table 9. Confusion Matrix CNN model

Model	Bidirectional LSTM Predict Class			GlobalMaxPooling Predict Class			LSTM Predict Class		
True Class	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive
Negative	212	52	36	221	54	25	222	42	36
Neutral	30	369	90	38	401	50	45	368	76
Positive	13	96	518	20	117	490	19	98	510

The results of calculating accuracy, precision, recall, and f1 score using the confusion matrix in CNN model are shown in Table 11.

At this optimization stage, overfitting and optimizing of the learning rate using the Callback Learning Rate Scheduler with a learning rate declaration of 10^{-7} were prevented, and then each iteration was increased up to 100 times iteration for each CNN model. The following Table 10 shows the result of the optimization learning rate for the CNN model.

Table 10. Optimization of Learning Rate Result

Model	***Best Loss***	***Best Learning Rate***
CNN - Bidirectional LSTM	0.11912	**0.00398**
CNN GlobalMaxPooling	0.11083	**0.00185**
CNN – LSTM	0.09894	**0.00185**

The following is an example of a graph at the learning rate optimization stage as can be seen in Fig.3, Fig.4 and Fig.5.

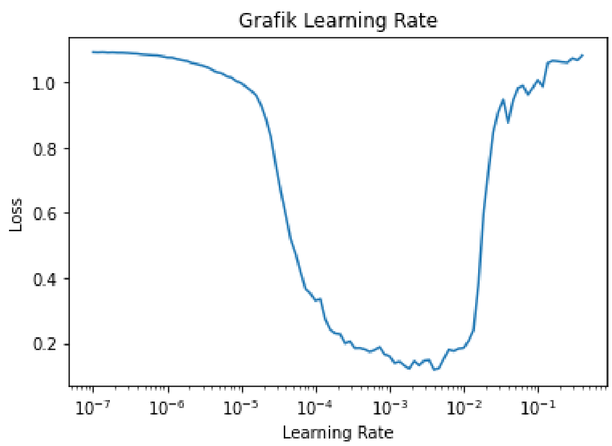

Figure 3. CNN-Bidirectional LSTM learning rate graph

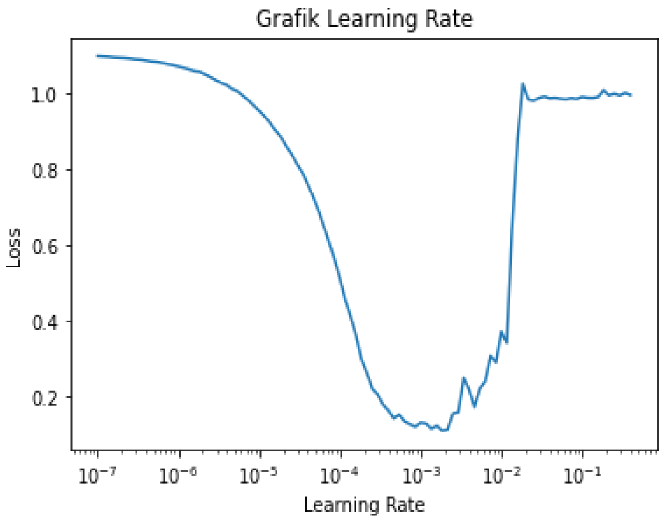

Figure 4. CNN GlobalMaxPooling learning rate graph

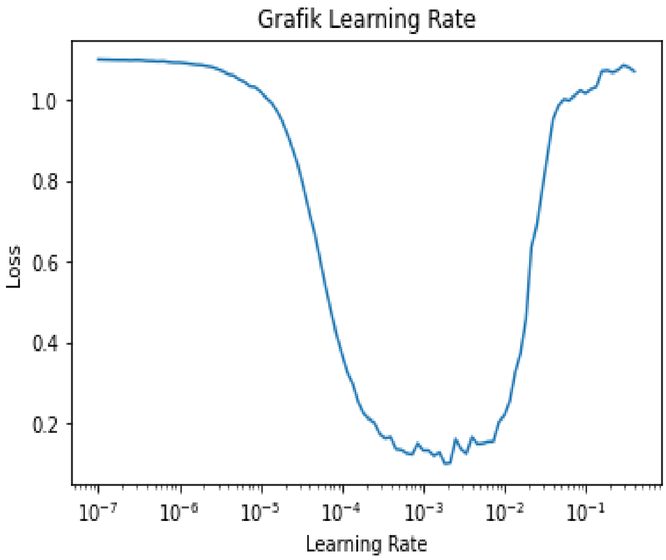

Figure 5. CNN-LSTM learning rate graph

3.5 Evaluation

This study used split and k-fold validation techniques with a confusion matrix. The split and k-fold validation function was the accuracy test of the training learning outcomes data, while the confusion matrix was information about the actual and predictions given by the classifier. This study used 4875 data which were then performed with 7:3 data splits. The results from the tested split data for all models are presented in Table 11.

Table 11. Performance Results of All Models

Model	*Precision*	*Recall*	*F1-Score*	*Accuracy*
NBC	0.79	0.74	0.75	**0.762**
CNN - Bidirectional LSTM	0.78	0.76	0.77	**0.776**
CNN GlobalMaxPooling	0.78	0.79	0.78	**0.785**
CNN – LSTM	0.77	0.77	0.77	**0.78**

This study also used k-fold cross-validation from the sci-kit-learn library to get the accuracy value of using split data with a value of k=5 and using Stratified K-fold. K-fold is a cross-validation method used to estimate the capabilities of a machine learning model on invisible data. It is usually used to validate the model because it is easy to understand and to implement, and the results can show a higher informative value than ordinary Validation Methods. In k-fold, the training data were separated into five parts randomly, then each part was used as training data. Test data were carried out in each experiment. Each part became test data once and became training data next time [16].

The following table shows the result of k-fold k=5 for all models. It reports that the best CNN model is CNN GlobalMaxPooling model with the accuracy of 81.9%, and the learning rate optimization gets the accuracy of 85.4%. CNN GlobalMaxPooling model was then used in the implementation phase with the NBC model, see Table 12.

Table 12. Results of k-fold k=5 for All Models

Model	*Learning Rate*	*Average accuracy*
NBC	-	**0.753**
CNN - Bidirectional LSTM	0.001	**0.799**
	0.00251	**0.822**
CNN GlobalMaxPooling	0.001	**0.819**
	0.00184	**0.854**
CNN – LSTM	0.001	**0.799**
	0.00185	**0.835**

3.6 Implementation

The classification process applied the Naive Bayes method and Convolutional Neural Network method on the tested data, preparing as many as 150 data. The implementation process used the best model with split data testing, with the highest accuracy results. In the NBC model, the classification results were 55% of positive sentiment, 30.5% of neutral sentiment, and 14.6% of negative sentiment. The classification results of the CNN model without learning rate optimization showed 53% of positive sentiment, 34.4% of neutral sentiment, and 12.6% of negative sentiment. The CNN model with learning rate optimization got 43.7% positive sentiment classification, 37.7% neutral sentiment, and 18.5% negative sentiment.

In the classification process, there were several wrong classifications. The researcher assumed that it was caused by unbalanced data, that there might also be irrelevant data that resulted in data bias in the training stage and could change the model's knowledge of the data, and that the model might be too simple. To solve some of these problems, the researcher suggested using the topic modeling method to find out what topics would be dominant in each sentiment, and thus the classification results needed to be further processed with topic modeling.

Table 13 shows an example of the prediction results from the test data.

Table 13. Example of Test Data Prediction Results

Original Tweet	Tweet after Preprocessing	CNN model prediction results	CNN model prediction results + Learning rate	NBC Prediction Results
Kartu Prakerja Dinilai Jawab Kebutuhan Pekerjaan Masa Depan https://t.co/U7Vy9LupyZ	kartu prakerja nilai butuh kerja	neutral	negative	positive
Program Kartu Prakerja bisa membantu pemerintah untuk memperbaiki keterampilan angkatan kerja dari sisi supply. #KartuPrakerja https://t.co/3xps3lV0o4	program kartu prakerja bantu perintah terampil angkat kerja sisi supply kartuprakerja	neutral	positive	positive
CORE: Kartu Prakerja Mampu Kantongi 387 Ribu Jadi Wirausaha - https://t.co/wBUsEfOT6Z	core kartu prakerja kantong ribu wirausaha	positive	positive	positive

Status 8.000 Peserta Kartu Prakerja Gelombang 16 Terancam Dicabut https://t.co/Sv5XVViUad https://t.co/ZHBIcQkC9I	status kartu prakerja gelombang ancam pergi	positive	negative	neutral
@ovo_id tolong bantu saya krn saya suda kesal sama ovo.masa saya mau tranfer uang prakerja sa ko susah sekali suda1 bulan saya tidak bsa tranfer uang saya.	tolong bantu suda kesal ovo masa tranfer uang prakerja sa ko susah suda tranfer uang	neutral	neutral	negative
THR Tanpa Tukin, Alasan Sri Mulyani: Dananya Buat Kartu Prakerja hingga BPUM https://t.co/EUObq9uWrC	thr tukin alas sri mulyani dana kartu prakerja bpum	neutral	negative	neutral
Program Kartu Prakerja Dapat Membantu Pemerintah Memulihkan Ekonomi Nasional https://t.co/Sa4YCW2upC via @harianbatampos	program kartu prakerja bantu perintah pulih ekonomi nasional via	positive	positive	positive
Thread~ Mempertanyakan Kesigapan perangkat pemerintah khususnya pada Empat Titik Rawan Korupsi Penanganan Covid-19 di Indonesia. Dari Kartu Prakerja, Bansos hingga Tenaga Kesehatan comment, retweet dan like yaa Kimia Farma #SekatAntisipasiMudik	thread sigap perangkat perintah khusus titik rawan korupsi tangan covid indonesia kartu prakerja bansos tenaga sehat comment retweet like iya kimia farma sekatantisipasimudik	neutral	neutral	positive

4 CONCLUSION

Based on the analysis, it can be concluded that this research has been able to achieve its stated goals, including the followings:

1. To implement sentiment analysis on Prakerja (pre-employment) Card program, this research applied three architectural models of Convolutional Neural Network: CNN - Bidirectional LSTM, CNN GlobalMaxPooling, and CNN - LSTM. Among the three, the best Convolutional Neural Network architecture is the CNN model with the MaxPooling layer, which results in 78.5% accuracy on Confusion Matrix and 81.9% on k-fold evaluation.
2. Sentiment analysis using the Naïve Bayes Classifier (NBC) method and the Convolutional Neural Network (CNN) method can classify sentiments about the Prakerja Card program in Indonesia. The accuracy when using CNN method with GlobalMaxPooling towards the keyword "Prakerja" found in 4725 tweets data reached 78.5%. NBC layers only got 76.2% accuracy on Confusion Matrix with a train and validation data separation ratio of 7:3. In evaluating the k-fold cross-validation model using StratifiedKFold with five classes, the best average accuracy result was 85.4% on the GlobalMaxPooling CNN layer model with a learning rate optimization of 0.00158. While the average accuracy on NBC only reached 75.3%.
3. Based on the implementation of sentiment analysis using the Naïve Bayes Classifier (NBC) method and the Convolutional Neural Network (CNN) method on the test data of 150 tweets, this research found that in the NBC model, the classification results for positive sentiment was 55%, neutral sentiment was 30.5%, and negative sentiment was 14.6%. The CNN model with learning rate optimization got 43.7% positive sentiment classification results, 37.7% neutral sentiment, and 18.5% negative sentiment. Finally, on the CNN model without learning rate optimization, the classification results for positive sentiment was 53%, neutral sentiment was 34.4%, and negative sentiment was 12.6%.

AUTHOR'S CONTRIBUTION

As the first author, Pahlevi Wahyu Hardjita contributed technically and wrote this research. He did the research with the supervision from the second author Nurochman for ideas and technical aspects. Last, the researcher got theoretical supervision from the third author, that is, Rahmat Hidayat.

COMPETING INTERESTS

Complying with the publication ethics of this journal, Pahlevi Wahyu Hardjita, Nurochman, and Rahmat Hidayat as the authors of this article declare that this article is free from conflict of interest (COI) or competing interest (CI).

ACKNOWLEDGMENT

Thank you to all parties for all input, moral and material support, corrections, and assistance throughout the research process so that this research can be completed.

REFERENCES

[1] C. Akbar, "Pandemi COVID-19, Airlangga: Pekerja Yang Di-PHK Naik Jadi 2,1 Juta Orang," *Tempo.co*, 2020.

[2] A. D. Handayani and A. Rachman, "Dampak Pandemi COVID-19 Terhadap Motif dan Tren Penelusuran Informasi Kartu Prakerja di Indonesia," *Empower J. Pengemb. Masy. Islam*, vol. 5, no. 2, p. 111, 2020, doi: 10.24235/empower.v5i2.7187.

[3] M. P. K. Prakerja, "Apa Itu Program Kartu Prakerja?," *Kementerian Koordinator Bidang Perekonomian Republik Indonesia*.

[4] F. Ronen and J. Sanger, *The Text Mining Handbook : Advanced Approaches in Analysing Unstructured Data*. New York, NY, USA: Cambridge University Press, 2007.

[5] Ii. Budiman, M. R. Faisal, and D. T. Nugrahadi, “Studi Ekstraksi Fitur Berbasis Vektor Word2Vec pada Pembentukan Fitur Berdimensi Rendah,” *J. Komputasi*, vol. 8, no. 1, pp. 62–69, 2020, doi: 10.23960/komputasi.v8i1.2517.

[6] M. Chalida and M. D. R. Wahyudi, “Analisis Sentimen Ujaran Kebencian Pemilihan Presiden 2019 Menggunakan Algoritma Naïve Bayes (Studi Kasus: Tweet #Pilpres2019 Di Kota Jakarta, Bandung, Semarang, Surabaya Dan Yogyakarta),” *Jnanaloka (Jurnal Open Access Yayasan Lentera Dua Indones.*, no. 2001, pp. 5–10, 2019.

[7] A. F. Hidayatullah and R. A. N. Nayoan, “Analisis Sentimen Berbasis Fitur Pada Ulasan Tempat Wisata Menggunakan Metode Convolutional Neural Network(CNN),” 2019.

[8] D. Prasetyawan and S. ’Uyun, “Penentuan Emosi Pada Video Dengan Convolutional Neural Network,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 5, no. 1, p. 23, 2020, doi: 10.14421/jiska.2020.51-04.

[9] M. H. Widianto, “Algoritma Naive Bayes,” *Binus Univesity*, 2019.

[10] J. Wang, L. C. Yu, K. R. Lai, and X. Zhang, “Dimensional sentiment analysis using a regional CNN-LSTM model,” *54th Annu. Meet. Assoc. Comput. Linguist. ACL 2016 - Short Pap.*, pp. 225–230, 2016, doi: 10.18653/v1/p16-2037.

[11] M. Theo and A. Bangsa, “Analisis Sentimen Berbasis Aspek Pada Ulasan Toko Online Menggunakan Convolutional Neural Network MHD. THEO ARI BANGSA, Dr. Yohanes Suyanto, M.I.Kom; Dr. Sigit Priyanta, S.Si, M.Kom,” pp. 84–85, 2019.

[12] A. Deviyanto and M. D. R. Wahyudi, “Penerapan Analisis Sentimen Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 3, no. 1, p. 1, 2018, doi: 10.14421/jiska.2018.31-01.

[13] B. Gunawan, H. S. Pratiwi, and E. E. Pratama, “Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes,” *J. Edukasi dan Penelit. Inform.*, vol. 4, no. 2, p. 113, 2018, doi: 10.26418/jp.v4i2.27526.

[14] J. A. Septian, T. M. Fahrudin, and A. Nugroho, “Analisis Sentimen Pengguna Twitter Terhadap Polemik Persepakbolaan Indonesia Menggunakan Pembobotan TF-IDF dan K-Nearest Neighbor,” *J. Intell. Syst. Comput.*, pp. 43–49, 2019.

[15] A. Hendra and F. Fitriyani, “Analisis Sentimen Review Halodoc Menggunakan Nai ̈ve Bayes Classifier,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 6, no. 2, pp. 78–89, 2021, doi: 10.14421/jiska.2021.6.2.78-89.

[16] A. R. Susanti, T. Djatna, and W. A. Kusuma, “Twitter’s sentiment analysis on GSM services using Multinomial Naïve Bayes,” *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 15, no. 3, pp. 1354–1361, 2017, doi: 10.12928/TELKOMNIKA.v15i3.4284.

[17] S. D. Harijiatno, “Analisis Sentimen Pada Twitter Menggunakan Multinominal Naive Bayes,” *J. Phys. A Math. Theor.*, vol. 8, no. 5, pp. 1–104, 2019.