Sentiment Analysis of the Use of Telecommunication Providers on Twitter Social Media using Convolutional Neural Network

Zafiratul Amalia
Department of Informatics
UIN Sunan Gunung Djati Bandung
Bandung, Indonesia
zafiratulamalia@gmail.com

Mohamad Irfan
Department of Informatics
UIN Sunan Gunung Djati Bandung
Bandung, Indonesia
irfan.bahaf@uinsgd.ac.id

Dian Sa'adillah Maylawati
Department of Informatics
UIN Sunan Gunung Djati Bandung
Bandung, Indonesia
diansm@uinsgd.ac.id
https://orcid.org/0000-0002-1193-3370

Agung Wahana
Department of Informatics
UIN Sunan Gunung Djati Bandung
Bandung, Indonesia
wahana.agung@uinsgd.ac.id

Wildan Budiawan Zulfikar
Department of Informatics
UIN Sunan Gunung Djati Bandung
Bandung, Indonesia
wildan.b@uinsgd.ac.id

Muhammad Ali Ramdhani
Department of Informatics
UIN Sunan Gunung Djati Bandung
Bandung, Indonesia
m_ali_ramdhani@uinsgd.ac.id
https://orcid.org/0000-0002-6492-067X

***Abstract*— Telecommunication technology continues to develop starting from 1G, 2G, 3G, 4G, and currently entering the 5G era. The Global System for Mobile Communications (GSM) based telecommunication industry in Indonesia consists of three big names: Telkomsel, XL, and Indosat. During the Covid-19 pandemic, activities carried out outside the home should be done online. People hope that the internet network can work properly. However, the reality is not as expected, because many networks are experiencing slow internet problems and many complaints are delivered through social media. Therefore, this research aims to find the insight opinions that have been conveyed to the telecommunications operator in social media. This research used the Convolutional Neural Network (CNN) algorithm to classify text sentiment (negative or positive) about telecommunication providers. The experiment with text data from Twitter is conducted after preprocessing and weighting of the Word2Vec process. The confusion matrix experiment shows that the CNN algorithm's performance reaches an average accuracy value of around 86.22%. The experiment was carried out by dividing the training data and testing the data 5 times in 10 times. The study results indicated that disruption of cellular telecommunications operators provided many sentiments, especially negative sentiment at the beginning of the COVID-19 pandemic.**

***Keywords*— convolutional neural network, deep learning, sentiment analysis, telecommunication provider**

I. Introduction

The development of communication technology in Indonesia has entered a new phase with the development of the information technology industry. Cellular phone coverage has reached all provinces in Indonesia with some regencies/cities in Indonesia. Telecommunication technology is also providing more types of telecommunications services, ranging from fixed telephones, wireless telephones, mobile telephones, etc. The role of the cellular telecommunications industry in people's lives and the national economy is crucial. The telecommunications sector tends to increase from 2017 to 2020 [1].

Mobile telecommunications technology on the internet network will continue to develop, starting from 1G, 2G, 3G, and until now entering the 4G era. However, several telecommunications companies are already preparing to enter the latest era, namely the 5G era. Along with the times, until now the GSM-based telecommunications industry occupies three big names in the telecommunications operator market in Indonesia, namely: Telkomsel, XL Axiata, and Indosat [2]–[4].

In 2017, 3 telecommunications operator companies provided internet services, which ranked in the top 3 based on a survey of mobile-based internet users. 43% of users use Telkomsel, 18.1% use Indosat, and 18% use XL. The main reason Indonesian people use mobile operators for mobile-based internet service providers is that the strongest signal at home gets 52.3% survey results, affordable package prices get 18.3% survey results, and the number of promos and bonuses with 8.5% of survey results [5]. Moreover, in 2021, Telkomsel, Indosat, and XL still be the biggest telecommunication provider in Indonesia [6], [7]. Then, Regulation of the Minister of Communication and Informatics No: 48/PER/M.KOMINFO/11/2009 concerning internet access service providers in the sub-district internet universal telecommunications service area to support internet access services for the community as well as to encourage the accelerated improvement of services and for the use of information and communication technology which has the aim of increasing the intelligence of citizens and the welfare of the community, a mobile internet network service is needed [8].

The Covid-19 pandemic has forced all Indonesians to carry out daily activities only at home. School children or students who are currently studying are affected by this pandemic. They are required to take online learning. Some office workers were also affected by requiring them to work online. The Indonesian people hope that the telecommunication operator signals they use will run smoothly without any problems from the following online activities. However, many Indonesians may complain about the telecommunication operator's network that is used frequently experiencing disturbances. Based on research released by Hootsuite, in January 2020, Indonesia's Internet speed had an average of only 20.1 Mbps or far below the world's average internet speed of 73.6 Mbps. In his research,

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

Singapore occupies a country with a record internet speed of up to 200.1 Mbps [9].

This condition gave rise to many opinions including on social media Twitter. Twitter is the most data source that used in text analytics research. This is because of the easy access and permission to pull data from Twitter compared to other social media. However, apart from more uncomplicated data collection techniques, Twitter is a place that is more open, open-minded, egalitarian, and a place for various trending news [10]. Compared to Facebook, which has many fake accounts and is spreading hoax news, and Instagram, which is mostly used for self-promotion and self-exposure, Twitter has far more educated users. The lack of mental judgment shows this, and if there is news that is a hoax, has a toxic smell, and is fighting with each other, Twitter users are wiser in dealing with it. In essence, social media analysis technology is currently one of the research strengths in the digital era [11]. Therefore, Twitter is considered to be more in demand by the public because it is easy to express their opinions. With so many Twitter users expressing this opinion, it can be used to find information. However, its use requires good analysis so that the information obtained can help many parties to support a decision.

Today, many sentiment analysis researchers use the Deep Learning method for Natural Language Processing (NLP), including sentiment analysis. The Deep Learning method is considered to have higher accuracy and is consistent when faced with big data. Deep Learning can produce improved accuracy as a result of these advantages [7]. There are Indonesia's largest NLP communities, such as IndoLEM, are developing Indonesian NLP benchmarks (Indonesian Language Evaluation Montage) [12], IndoNLU (Indonesian Natural Language Understanding) [13], and IndoNLG (Indonesian Language for Natural Language Generation) [14]. These studies also perform NLP research, which includes sentiment analysis utilizing Deep Learning. Deep learning, which is the creation of artificial neural networks, is widely used in sentiment analysis: (1) sentiment analysis research which proves sentiment analysis using CNN has a better accuracy value than using other neural network models, such as: Recursive Neural Network (RNN) and Matrix-Vector Neural Network (MV RNN). From the dataset that has been obtained, as many as 10,662 film reviews, half of which are positive data and half are negative data. The accuracy value obtained using the CNN algorithm reaches 80% [15]; (2) short text sentiment analysis to prove the effectiveness of the convolutional neural network algorithm compared to traditional methods in sentiment analysis of comments from users' attitudes based on various texts collected from Weibo [16]; (3) sentiment analysis of the Russian community for reviews of Russian-language e-commerce products located in Russia. In the test by including emoji, the f-measure result was 75.45% [17]; (4) compare the advantages of each algorithm, namely the CNN algorithm and the Support Vector Machine (SVM) in sentiment analysis. In this journal, a sentiment analysis model that uses CNN and SVM algorithms can work effectively and improve the performance of text classification [18], [19]; (5) Sentiment analysis for ranking the popularity of tourist destinations which aims to find out popular tourist attractions according to social media users based on the results of likes, statuses, and accounts that discuss these tourist attractions [20]; and so on.

From the ongoing problems regarding the speed of the internet network in Indonesia, which is experiencing disruption or slowness, Indonesia is still stuck on 3G and 4G networks even though several countries have started accessing 5G networks. With the analysis of data from Twitter regarding the use of telecommunication operators. It is hoped that it can provide information to operator users to be more selective in choosing and help internet service providers improve their internet network and as a benchmark whether it is feasible to enter the 5G internet network. Therefore, this study aims to analyze the public sentiment about the use of telecommunication providers through Twitter using Convolutional Neural Networks.

II. Research Methods

A. Sentiment Analysis

Sentiment analysis is a method for determining public opinion on certain topics [21], [22]. Sentiment analysis classifies data with positive, neutral, or negative class labels relating to public opinion or opinions from a variety of sources, including social media [23]–[28], movie reviews [29]–[31], news portals [32], product reviews [33], [34], and others. The sentiment analysis technique employs a classification strategy. The training data must be labeled (positive, negative, or neutral) for the computer to learn to recognize patterns and predict labels on new text data.

B. Research Activities

This research was conducted with a data science approach using the data science methodology CRISP-DM [35], [36](Cross-Industry Standard Process for Data Mining) (P Chapman et al., 2000; Pete Chapman et al., 1999). Activities in the CRISP-DM methodology consist of business understanding, data understanding, data preparation, modeling, model evaluation, deployment, and model management. Where, the activities contained in CRISP-DM are listed in Fig. 1 in detail, including:

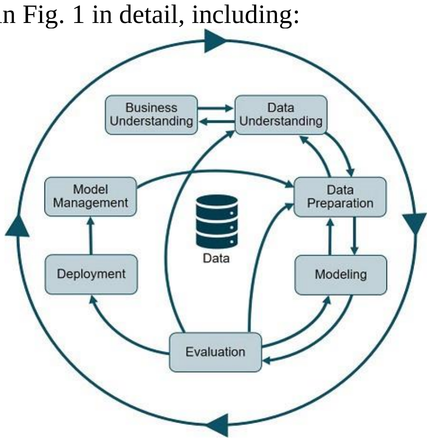

Fig. 1. CRISP-DM Methodology [37]

1. Business Understanding is an activity to identify and explore organizational needs, determine business objectives, determine data science technical goals, and plan data science projects. At the business understanding stage, an analysis of what approach or technical data science will be used is also carried out in accordance with the business problem to be solved.

2. Data Understanding is the process of understanding and reviewing data requirements that will be used to solve business problems that have been analyzed. In this data understanding activity, data collection is carried out for further review and validation.
3. Data preparation is the initial process before developing the model. Activities carried out include sorting, cleaning, constructing, integrating, and labeling data. This process also produces data visualization that facilitates data analysis.
4. Modeling is the main process in developing models used in solving data science problems. At this modeling stage, apart from building the required model, we also develop test scenarios for the model.
5. Evaluation is a process after several model testing scenarios have been carried out. This evaluation is carried out to select the best model and ensure that the model is well implemented and successfully resolves business problems.
6. Deployment is the process of implementing a model in the form of an application or software that can be accessed by end-users¬. At this stage, the activities carried out include making a model deployment plan, carrying out the model deployment process, making a maintenance plan, and performing maintenance on the model and application.
7. Model Management can contain the management and maintenance of the model by getting feedback. Feedback is the process of evaluating the entire data science process that has been carried out by reviewing and making data science output reports.

This research conducts the business understanding until the evaluation of the model. The deployment and model management are not conducted.

C. Convolutional Neural Network

Convolutional Neural Networks (CNN) is a type of neural network that is widely utilized in image and text processing. Convolution, or simply convolution, is a filtering and classification matrix for images and text [41]. Filtering is done in each phase using the layers of a Convolutional Neural Network. The training procedure is the name of the technique. The training procedure is divided into three layers: Convolutional, Pooling, and Fully-connected layers [42][43]. The CNN layer architecture is shown in Figure 2. The Convolutional layer is the main layer for generating new features from input data. Like the Convolutional Layer, the Pooling layer is responsible for shrinking the Convolved Feature's spatial size. Because of the dimensionality reduction, the amount of computing power required to process the data is reduced. It also helps with training by extracting rotational and positional invariant dominant characteristics. Pooling can be divided into two types: maximal pooling and average pooling. The maximum value from the image's Kernel-covered area is returned by Max Pooling. Average Pooling returns the average of all values from the Kernel's data segment. After that, the Fully-Connected layer starts learning a non-linear function. Learning non-linear combinations of the high-level qualities represented by the convolutional layer's output using a Fully-Connected layer is a cheap strategy for learning non-linear combinations of the high-level qualities represented by the convolutional layer's output.

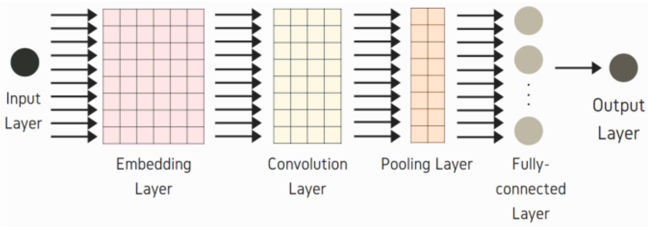

Fig. 2. CNN Architecture Illustration

III. Result and Discussion

A. Business Understanding

Based on research that has been done, internet speed in Indonesia has a meager average compared to several countries in Asia. Indonesia is even still trapped in 3G and 4G networks even though several countries have started accessing networks to 5G, and some have even entered the 6G network. If only the existing networks were more sophisticated, the costs would be even lower. The uneven internet service infrastructure is one of the causes of the current slow internet network. Development is only concentrated in big cities, while it is challenging to get access in remote areas and has not been served.

So far, telecommunications operators are competing for positions in areas that are crowded with users. This business pattern encourages telecommunications operators to develop their internet network infrastructure in regions with many uses, such as offices and shopping centers. Meanwhile, the development of the internet network in residential areas has received less attention because, in terms of business, the economy of scale in residential areas is low. This resulted in residential areas experiencing slow access to the internet network, which is widely felt today. With the existence of Large-Scale Social Restrictions (PSBB) and the large number of people doing Work from Home (WFH) in various regions, this has resulted in high community activities from home. To work from home, people also need internet access which is as fast as internet access when they work in the office.

In the use of comment data in the form of text tweets from social media, Twitter is often used to process each comment's sentiment in the form of tweets. Sentiment analysis system from users of telecommunication operators for internet network services becomes a means to express user satisfaction and complaints from telecommunication operators providing internet services. Therefore, in overcoming and knowing the problems of complaints that often occur to telecommunications operators, a system that can analyze social media users' opinions, especially on Twitter social media, is proposed.

B. Data Understanding

With the popular use of social media for people in Indonesia, they are indirectly connected to each other and exchange information that they know. Freedom in its use also gives rise to many opinions or opinions that make them active in responding to social issues that are happening both at home and abroad. The data taken for this sentiment analysis research is taken by utilizing the Twitter API provided by

Twitter for free so that data retrieval can be done more efficiently than other types of social media.

On social media, Twitter is the official account owned by an internet service provider that is used to interact with customers or their users. The data taken is in the form of tweets that are directly mentioned to the official accounts @Telkomsel, @IndosatCare, and @myXLCare. Tweet data is taken starting from May 15, 2020 to August 10, 2020. The number of users taken as the example above for the number of mentions on official accounts for the period May 16, 2020 - May 31, 2020, namely: a total of 491 mentions @Telkomsel, 407 mentions @IndosatCare, and 318 mentions of @myXLCare. The data will be grouped or classified into two sentiments, there are positive sentiments and negative sentiments. A lot of tweet data is in the form of criticism or complaints as well as the satisfaction felt with this sentiment analysis system. That's why only positive and negative sentiments are taken, and no neutral sentiments are included.

From the dataset that has been collected, it will be used as training data (training data) and testing data (test data). Experiments on the system will be carried out five times with variations in training and testing data used. The experiment was carried out successively using 90% training data and 10% testing data, 80% training data and 20% testing data, 70% training data and 30% testing data, 60% training data and 40% testing data, and 50% training data and 50% of testing data.

C. Data Preparation

The data taken from Twitter is in raw or unstructured data, which causes the data to be processed first. Unstructured data is not well used in data mining because the data still contains a lot of noise, possibly incomplete data. Noise is a very sensitive thing in data mining. Data needs to be processed or improved so that the resulting output becomes higher quality data, and of course, the data will be cleaner, namely, text preprocessing.

Text preprocessing is an essential step in the data preparation stage of the mining process. Text preprocessing guarantees that input data is of high quality, guaranteeing that the outcomes are as expected [38], [39]. Case-folding and tokenizing, deleting superfluous letters, removing stop-words, and stemming are among the text preprocessing techniques used in this study. The Sastrawi library is used for Indonesian language stemming, with the Nazief-Adriani stemming method [40]. Table I shows the example process of text preprocessing (begin from case-folding, filtering, stop-words removing, and stemming) from the example of three tweets in the Indonesian language below:

1. *SETIAP HUJAN SINYAL LANGSUNG JELEK AJA! @Telkomsel tolong dong diperbaiki. Bayar mahal masa buat kaya gini doang?*(EVERY RAIN SIGNAL IS BAD IMMEDIATELY! @Telkomsel please fix it. Do you pay dearly for something like this?)
2. *Woi.... @IndosatCare ,,,sinyal datamu kemana kok ngilang mulu,,,,* (Wow.... @IndosatCare ,,, where did your data signal disappear?)
3. *Ternyata 21gb ku juga dikasih bonus unlimited,, thx loh @myXLCare* (It turns out that my 21gb was also given an unlimited bonus,, thx, @myXLCare)

TABLE I. THE EXAMPLE OF TEXT PRE-PROCESSING RESULT

Process	Result
Case-folding	1. setiap hujan sinyal langsung jelek aja! @telkomsel tolong diperbaiki. bayar mahal masa buat kaya gini doang? 2. woi…. @indosatcare ,,, sinyal datamu kemana kok ngilang mulu,,,, 3. sinyal jelek banget, tolong dong ditolongin @myxlcare
Removing Regular Expression	1. setiap hujan sinyal langsung jelek aja telkomsel tolong diperbaiki bayar mahal masa buat kaya gini doang 2. woi indosatcare sinyal datamu kemana kok ngilang mulu 3. sinyal jelek banget tolong dong ditolongin myxlcare
Stop-words Removing	1. hujan sinyal jelek telkomsel tolong diperbaiki bayar mahal buat gini 2. indosatcare sinyal data kemana ngilang 3. sinyal jelek tolong myxlcare
Stemming	1. hujan sinyal jelek telkomsel tolong perbaiki bayar mahal buat gini 2. indosatcare sinyal data kemana hilang 3. sinyal jelek tolong myxlcare

D. Word Embedding using Word2Vec

Word embedding is a type of word representation that allows for representing words that have similar meanings [48]. There are many word embedding representations, such as Word2Vec [49], GloVe [50], and FastText [51]. The embedding layer in this study is Word2Vec. Word2vec is a method for expressing words in N-dimensional vector form [52]. Word2Vec uses a neural network to determine contextual and semantic similarity (contextual and semantic similarity) in the form of one-hot encoded vectors when presenting a word. Contextual and semantic similarity can be used to illustrate a word's relationship to other words. Word embedding representation varies by language.

Word weighting (Word2Vec) is done after going through the data preprocessing process. In Word2Vec words will be represented in the form of vectors and create a vocabulary with models from training data. These vocabulary words will form vectors that place words, so words with the same meaning will be close together in the vector space. Vector values that are owned by almost the same word will have almost the same value. Word2vec has various dimensions, there are 300, 250, 200, 150, 100, and 50. It would be easier if the analyzed data had the same size. To adjust the length and short of a sentence will be added a space to fill the empty matrix. After generating the matrix of the same length, the data is ready to be used in the following process, namely as input to the model process of the convolutional neural network algorithm that will be used. Table II shows the example of Word2Vec from the first tweet in Table I.

TABLE II. THE EXAMPLE OF WORD2VEC RESULT

Words	Result
'hujan'	[-1.169715, 0.704697, 1.31705 , -0.110299, -0.148704, -1.2512 ,-0.875249, -0.320313, -0.858832, 0.804166, 1.098714, -0.17406 ,0.482215, 0.338141, -1.650685, 0.797885, -1.012358, 0.92607 ,-0.1301 19, 0.838666, -1.468397, 0.397198, 1.185056, 0.202073,0.275191, -0.450439, 1.462456, 0.060101, 0.756593, 0.42172 ,0.663162, 1.059709, -0.746781 , 0.17531 , 0.644469, 1.10234 ,1.407392, 0.878 321, 1.747291, -0.400864, 0.303235, -0.731086,-0.213796, -1.893042, -0.608188, 0.649766, -0.57593 4, -0.239684,0.414374, 0.65975]
…	…
'mahal'	[-0.542085, 0.932004, 0.298914, -0.682565, -0.320728, -1.991384,-0.568624, 0.109093, -0.032789, 0

	.151607, -0.088974, 0.910611,-0.303522, 0.333293 , 0.027298, 0.971275, 1.137367, 0.797534,-0.45 1703, 0.455621, -0.939714, -0.437153, 0.531893, 0.488844,0.233192, -0.648929, -0.811408, 1.049156 , 0.465909, 0.226404,-0.408238, 0.601962, 0.40 0628, 1.359354, 0.264591, 0.181183,0.459421, 0 .149649, 0.032209, 0.388646, 0.349456, -0.74412 6,0.857827, -0.00797 , 0.788202, 0.397383, -0.43 1107, -0.711504,0.258034, -0.234832]

E. Modeling Process with Convolutional Neural Network

By using the linear regression layer as the output layer, we can create a simple CNN model for our purpose. The task's CNN architecture is given here. The embedding layer, which is the initial layer of the model, is found within the CNN algorithm. Words are represented as real-valued vectors in a high-dimensional space using layer embedding. This layer allows for the pre-trained word vector matrix to be used to initialize vocabulary word vectors. The input tweets are converted into numeric word token sequences, such as t_1, t_2, ..., t_N, where t_N represents the original word, and N is the token vector length. To keep the result size identical for tweets of varying length, the authors limit the maximum value of N to the maximum tweet length of all tweets. Any tweet was shorter than N will be populated to N using zeros. In the convolutional layer, the m filter is used to extract the local n-gram features from the matrix of the previous layer.

There are several stages in the classification process, namely Convolutional Layer, Pooling Layer, and Fully Connected Layer. After the preprocessing stage, the module's word (tweet) input will be converted into an array of tokens with specific dimensions using word embedding and then mapped in the feature matrix or sentence matrix by the embedding layer. Furthermore, in the data convolution layer, the weight is calculated layer by layer with a filter layer to perform calculations on each embedding layer. This layer is used to form the weights that the input layer has with the hidden layer. Then, the largest value weight is taken from each filter layer using a pooling layer. The fully connected layer finally composes these features to output the final regression results with a linear decoder. After doing the pre-training then it is done to train the backpropagation algorithm to train the error rate. A loss function is an error in the training data set. Loss validation is an error after running a data validation set.

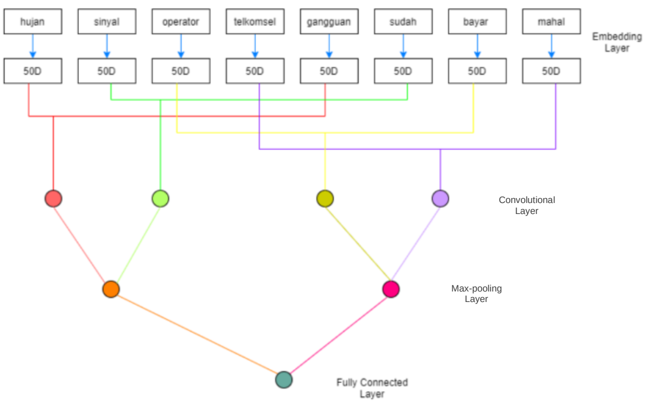

Fig. 3. Illustration of CNN Modeling Process

TABLE III. THE EXAMPLE OF CNN MODELING PROCESS

Embedding Process			
Words:	'hujan'	…	'mahal'
Embd. 1	-1.6	…	-0.54
Embd. 2	0.7	…	0.93
Embd. 3	1.31	…	0.29
Convolutional Process			
Convolutional 1	-0.462	0.628	0.686
Convolutional 2	-0.514	0.892	0.444
Convolutional 3	-0.42	0.912	0.492
Convolutional 4	-0.46	0.978	0.426
Max-pooling Process			
Max-pooling 1	-0.462	0.892	0.686
Max-pooling 2	-0.42	0.978	0.492
Fully Connected Process			
0.978			

Fig 3 and Table III illustrate the CNN modeling process. At the pooling layer, the unification operation will be carried out which is used to combine vectors generated from different convolution layers into one-dimensional vectors. This process is done by taking the maximum value observed in the vector generated from the convolution layer. This vector will later capture the features that most relate to the text. From the illustration above, one of the examples of tweets used in this research is used. Each word has a weighting in the form of a 50x50 matrix, but in the illustration above, it only takes 3 matrices.

Furthermore, in the convolution layer, the number of weighting results is calculated and then divided by five because the coding stores the kernel which is worth 5. After getting the results from the calculations in the convolution layer, the largest value for each row is sought from convolution one and convolution two columns and the convolution 3 and 4 columns. The comparison results will enter the max-pooling layer, find the largest value, and end up with the final connected delayer. The value in the final connected layer is the largest value after comparing it to the max-pooling layer.

E. Evaluation

Evaluation is carried out to adjust whether the results of the modeling with the desired business understanding are appropriate or not. Before the evaluation is carried out, it will be tested first. By testing, we can find out whether the system built can achieve the desired business understanding. Twenty tests using epoch 50 and distribution between each official account @Telkomsel, @IndosatCare, and @myXLCare, with 90% sharing of data as training data with 10% testing data, 80% training data with testing data as much as 20%, training data as much as 70% with 30% testing data, 60% training data with 40% testing data, 50% training data with 50% testing data.

This research used the confusion matrix to evaluate the sentiment classification result model [44]. There are values of accuracy, precision, and Recall, which are calculated by the formula (1)-(3) [45]–[47].

$$\text{precision} = \frac{|TP|}{|TP|+|TN|} \tag{1}$$

$$\text{recall} = \frac{|TP|}{|TP|+|FP|} \tag{2}$$

$$\text{accuracy} = \frac{|TP|+|FN|}{|D|} \tag{3}$$

Precision is calculated by dividing the total sentiment classifications that are correctly predicted to be positive (True Positive - TP) by the number of sentiment classifications that are correctly predicted to be positive and the correct sentiment classifications are predicted to be negative in the document (True Negative - TN), so that the prediction results correctly determine whether the sentiment classification is correct, both predicted by humans and predicted by the system is correct. In a document, Recall is equal to the total number of predicted true sentiment classifications divided by the number of false-positive predicted and false-positive sentiment classifications (FP). Recall evaluates the proportion of sentiment classifications predicted by humans and those generated by the system. At the same time, accuracy is defined as the proportion of the total sentiment classification that is predicted correctly and incorrectly divided by the total sentiment classification in a document.

F. Experiment Scenario and Result

There are 20 times experiment scenarios with 1,480 tweets data. Experiment I was conducted for the entire dataset consisting of mentions of official accounts @Telkomsel, @IndosatCare, and @myXLCare was carried out with 90% training data with 1,332 data and 10% testing data with 148 data. Experiment II for the entire dataset consisting of mentions of official accounts @Telkomsel, @IndosatCare, and @myXLCare was carried out with 80% training data with 1,184 data and 20% testing data with 296 data. Experiment III for the overall dataset consisting of mentions of official accounts @Telkomsel, @IndosatCare, and @myXLCare was carried out with 70% training data with 1,036 data and 30% testing data with 444 data. Experiment IV for the entire dataset consisting of mentions of official accounts @Telkomsel, @IndosatCare, and @myXLCare was carried out with 60% training data with 888 data and 40% testing data with 592 data. Experiment V test for the entire dataset consisting of mentions of official accounts @Telkomsel, @IndosatCare, and @myXLCare was carried out with 50% training data with 740 data and 50% testing data with 740 data. Experiment VI, to mention to the official account @Telkomsel was carried out with 90% training data with 452 data and 10% testing data with 50 data. Experiment VII to mention to the official account @Telkomsel was carried out with 80% training data with 402 data and 20% testing data with 100 data. Experiment VIII test to mention to the official account @Telkomsel was carried out with 70% training data with 351 data and 30% testing data with 151 data. Experiment IX to mention to the official account @Telkomsel was carried out with 60% training data with 301 data and 40% testing data with 201 data. Experiment X to mention to the official account @Telkomsel is carried out with 50% training data with 251 data and 50% testing data with 251 data.

Experiment XI to mention to the official account @IndosatCare was carried out with 90% training data with 438 data and 10% testing data with 49 data. Experiment XII to mention to the official account @IndosatCare was carried out with 80% training data with 390 data and 20% testing data with 97 data. Experiment XIII to mention to the official account @IndosatCare was carried out with 70% training data with 341 data and 30% testing data with 146 data. Experiment XIV test to mention to the official account @IndosatCare was carried out with 60% training data with 292 data and 40% testing data with 195 data. Experiment XV to mention to the official account @IndosatCare was carried out with 50% training data with 244 data and 50% testing data with 243 data. Experiment XVI for mentioning the official account @myXLCare was carried out with 90% training data with 442 data and 10% testing data with 49 data. Experiment XVII to mention the official account @myXLCare was carried out with 80% training data with 393 data and 20% testing data with 98 data. Experiment XVIII to mention the official account @myXLCare was carried out with 70% training data with 344 data and 30% testing data with 147 data. Experiment XIX test to mention to the official account @myXLCare was carried out with 60% training data with 295 data and 40% testing data with 196 data. Last, experiment XX to mention to the official account @myXLCare was carried out with 50% training data with 246 data and 50% testing data with 245 data. Table IV shows the accuracy result of those experiment.

TABLE IV. ACCURACY RESULT OF EXPERIMENT

Experiment Result of Whole of Data									
Splitting Percentage	Training Data	Testing Data	Precision (%)		Recall (%)		F1-Score (%)		Accuracy (%)
			(-)	(+)	(-)	(+)	(-)	(+)	
90%-10%	1332	148	95	88	98	79	96	83	93.59
80%-20%	1184	296	95	79	97	71	96	75	93.24
70%-30%	1036	444	93	79	97	63	95	70	91.44
60%-40%	888	592	94	83	98	64	96	72	92.39
50%-50%	740	740	95	72	95	75	95	74	91.48
Average Value			94.4	80.2	97	70.4	95.6	74.8	92.428
Experiment Result of @Telkomsel Official Account									
Splitting Percentage	Training Data	Testing Data	Precision (%)		Recall (%)		F1-Score (%)		Accuracy (%)
			(-)	(+)	(-)	(+)	(-)	(+)	
90%-10%	452	50	89	93	97	76	93	84	90.19
80%-20%	402	100	86	96	98	72	92	83	89.1
70%-30%	351	151	87	95	98	71	92	81	88.74
60%-40%	301	201	85	98	99	66	92	79	88.11
50%-50%	251	251	81	95	99	51	89	66	83.73
Average Value			85.6	95.4	98.2	67.2	91.6	78.6	87.974
Experiment Result of @IndosatCare Official Account									
Splitting Percentage	Training Data	Testing Data	Precision (%)		Recall (%)		F1-Score (%)		Accuracy (%)
			(-)	(+)	(-)	(+)	(-)	(+)	
90%-10%	438	49	97	100	100	25	97	40	93.87
80%-20%	390	97	96	100	100	33	98	50	95.91
70%-30%	341	146	93	0	100	0	96	0	92.66
60%-40%	292	195	93	0	100	0	96	0	92.83
50%-50%	244	243	95	100	100	13	97	24	94.67
Average Value			94.8	60	100	14.2	96.8	22.8	93.988

Experiment Result of @myXLCare Official Account									
Splitting Percentage	Training Data	Testing Data	Precision (%)		Recall (%)		F1-Score (%)		Accuracy (%)
			(-)	(+)	(-)	(+)	(-)	(+)	
90%-10%	442	49	85	50	98	12	91	20	83.99
80%-20%	393	98	88	25	97	8	92	12	85.85
70%-30%	344	147	91	100	100	13	95	24	91.21
60%-40%	295	196	93	80	99	24	96	36	92.89
50%-50%	246	245	91	80	100	16	95	27	91.05
Average Value			89.6	67	98.8	14.6	93.8	23.8	88.998

G. Discussion

In Table IV, the predictions generated from the convolutional neural network algorithm depend on the amount of data being trained. This is because the accuracy value of each test will be different with the use of different testing data. After testing 5 times on all datasets, mentioning the official accounts @Telkomsel, @IndosatCare, and @myXLCare, the average test results were 92.42% for testing all datasets; 87.97 for testing the @Telkomsel mention; 93.98% for testing the @IndosatCare mention; and 88.99% for testing the @myXLCare mention using epoch 50.

This experiment shows that the system is not running well. This is due to various factors. The first factor is that less data is used for this deep learning method. The data used is only 1480 tweets, this is due to the limited data collection time. This is because when crawling data through the tweet archive feature on the google spreadsheet, it can only retrieve 100 tweets per crawl. The next factor is the data used between negative data and positive data is not balanced. In the tests carried out, it was shown that the system was negative when classifying test tweets because 92% of the total tweet data that mentions the official accounts @IndosatCare and @myXLCare was labeled negative the system was learning it could not represent data that was labeled positive. Unlike the mention of @Telkomsel which has a balanced amount of positive and negative data.

Factors that affect the positive precision in the test are how many true negative and false positive values, while the negative precision is how many true negative values and false negative values. Furthermore, the factors that affect positive Recall are how many true positive and false negative values, while negative Recall is how many true negative values and false positive values. From the results of the labeling that has been done, it is found that negative sentiments have more numbers than positive sentiments. Prediction of negative sentiment is more accurate than positive sentiment because more negative data is tested than positive data.

CONCLUSION

The implementation of the sentiment analysis system using the Convolutional Neural Network algorithm has been successfully carried out by categorizing positive sentiment and negative sentiment. When the test was carried out five times for each data collection on mentions with official accounts @Telkomsel, @IndosatCare, and @myXLCare, the average test results were 87.97 for tests on mentions @Telkomsel; 93.98% for testing the @IndosatCare mention; and 88.99% for testing the @myXLCare mention using epoch 50. Several factors affect the level of accuracy: the first factor is that less data is used for this deep learning method. The data used is only 1480 tweets. This is due to the limited data collection time. The next factor is the data used between negative data and positive data is not balanced. In the tests carried out, the system when classifying test tweets became negative, because 92% of the tweet data that mentions the official accounts @IndosatCare and @myXLCare was labeled negative, so that when the system did the learning it could not represent data that was labeled positive. So it can be concluded that the telecommunications operator internet network services have negative sentiments from Twitter users. For future works, it needs more data collection is needed with a wider range of retrieval time. This is to improve the accuracy of sentiment analysis results.

REFERENCES

[1] Badan Pusat Statistik Indonesia, "Statistik Telekomunikasi Indonesia 2020," Jakarta, 2020. [Online]. Available: https://www.bps.go.id/publication/download.html?nrbvfeve=ZTAzYWNhMWU2YWU5MzM5NmVlNjYwMzI4&xzmn=aHR0cHM6Ly93d3cuYnBzLmdvLmlkL3B1YmxpY2F0aW9uLzIwMjEvMTAvMTEvZTAzYWNhMWU2YWU5MzM5NmVlNjYwMzI4L3N0YXRpc3Rpay10ZWxla29tdW5pa2FzaS1pbmRvbmVzaWEtMjAyMC5odG1s&twoadfnoarfeauf=MjAyMS0xMS0yOSAxNjoxOTo1Ng%3D%3D.

[2] K. Ariansyah, "Proyeksi Jumlah Pelanggan Telepon Bergerak Seluler di Indonesia," *Bul. Pos Dan Telekomun.*, vol. 12, no. 2, pp. 151–166, 2014.

[3] KOMINFO, "Menkominfo: 5 Tahun Lagi Hanya Ada 4 Operator Seluler," *kominfo.go.id*, 2021. https://kominfo.go.id/index.php/content/detail/4717/Menkominfo%3A+5+Tahun+Lagi+Hanya+Ada+4+Operator+Seluler/0/sorotan_media (accessed Nov. 01, 2021).

[4] N. Ghaidazia, "Ini Operator Seluler Terbaik di Indonesia Saat Ini, Download dan Upload Paling Ngebut!," *jalantikus.com*, 2021. https://jalantikus.com/telko/provider-tercepat-indonesia/ (accessed Nov. 01, 2021).

[5] Asosiasi Penyelenggara Jasa Internet Indonesia, "Penetrasi & Profil Perilaku Pengguna Internet Indonesia," *apjii.or.id*, 2017. https://apjii.or.id/survei2018s/download/TK5oJYBSyd8iqHA2eCh4FsGELm3ubj. (accessed Feb. 10, 2021).

[6] C. Stephanie, "Urutan Operator Seluler Terbaik di Indonesia untuk Streaming Video," *Kompas.com*, 2021. https://tekno.kompas.com/read/2021/07/01/14542067/urutan-operator-seluler-terbaik-di-indonesia-untuk-streaming-video?page=all (accessed Nov. 02, 2021).

[7] A. T. Haryanto, "Survei: Operator Seluler Terbaik Indonesia Adalah...," *detikInet*, 2021. https://inet.detik.com/telecommunication/d-5373152/survei-operator-seluler-terbaik-indonesia-adalah (accessed Nov. 02, 2021).

[8] Kementerian Komunikasi dan Informatika, "Peraturan Menteri Komunikasi dan Informatika Nomor 48/PER/M.KOMINFO/11/2009 tentang Penyediaan Jasa Akses Internet pada Wilayah Pelayanan Universal Telekomunikasi Internet Kecamatan," *kominfo.go.id*, 2009. https://jdih.kominfo.go.id/produk_hukum/view/id/406/t/peraturan+menteri+komunikasi+dan+informatika+nomor+48permkominfo112009+tanggal+23+november+2009.

[9] S. M. Jannah, "Di Balik Keluhan Konsumen soal Lemotnya Internet Selama WFH," *tirto.id*, 2020. https://tirto.id/di-balik-keluhan-konsumen-soal-lemotnya-internet-selama-wfh-forA (accessed Nov. 01, 2021).

[10] "Ciri Khas Yang Membedakan Pengguna Twitter, Instagram dan Facebook," *dipa14.web.id*, 2020. https://dipa14.web.id/2020/12/16/ciri-khas-yang-membedakan-pengguna-twitter-instagram-dan-facebook/ (accessed Aug. 11, 2021).

[11] A. I. Kabir, R. Karim, S. Newaz, and M. I. Hossain, "The Power of Social Media Analytics: Text Analytics Based on Sentiment Analysis and Word Clouds on R.," *Inform. Econ.*, vol. 22, no. 1, 2018.

[12] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP," *arXiv Prepr. arXiv2011.00677*, 2020.
[13] B. Wilie *et al.*, "IndoNLU: Benchmark and resources for evaluating Indonesian natural language understanding," *arXiv Prepr. arXiv2009.05387*, 2020.
[14] S. Cahyawijaya *et al.*, "IndoNLG: Benchmark and Resources for Evaluating Indonesian Natural Language Generation," *arXiv Prepr. arXiv2104.08200*, 2021.
[15] X. Ouyang, P. Zhou, C. H. Li, and L. Liu, "Sentiment analysis using convolutional neural network," in *2015 IEEE international conference on computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing*, 2015, pp. 2359–2364.
[16] W. Li, Z. Li, and X. Fang, "Short text sentiment analysis based on convolutional neural network," in *2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2018, pp. 291–295.
[17] S. Smetanin and M. Komarov, "Sentiment analysis of product reviews in Russian using convolutional neural networks," in *2019 IEEE 21st Conference on Business Informatics (CBI)*, 2019, vol. 1, pp. 482–486.
[18] K. Jain and S. Kaushal, "A comparative study of machine learning and deep learning techniques for sentiment analysis," in *2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, 2018, pp. 483–487.
[19] Y. Chen and Z. Zhang, "Research on text sentiment analysis based on CNNs and SVM," in *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2018, pp. 2731–2734.
[20] M. Murnawan, "Pemanfaatan Analisis Sentimen Untuk Pemeringkatan Popularitas Tujuan Wisata," *J. Penelit. Pos dan Inform.*, vol. 7, no. 2, pp. 109–120, 2017.
[21] B. Lin, F. Zampetti, M. Di Penta, R. Oliveto, G. Bavota, and M. Lanza, "Sentiment Analysis for Software Engineering: How Far Can We Go?," *2018 ACM/IEEE 40th Int. Conf. Softw. Eng. 117*, 2018, doi: 10.1145/3180155.3180195.
[22] B. Liu, "Sentiment analysis and opinion mining," *Synth. Lect. Hum. Lang. Technol.*, vol. 5, no. 1, pp. 1–184, 2012, doi: 10.2200/S00416ED1V01Y201204HLT016.
[23] A. P. Rodrigues, A. Rao, and N. N. Chiplunkar, "Sentiment Analysis of Real Time Twitter Data Using Big Data Approach," *2nd Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solut. CSITSS 2017*, pp. 1–6, 2018, doi: 10.1109/CSITSS.2017.8447656.
[24] A. Ali, "Sentiment Analysis on Twitter Data using KNN and SVM," vol. 8, no. 6, pp. 19–25, 2017.
[25] S. Rosenthal, N. Farra, and P. Nakov, "SemEval-2017 Task 4: Sentiment Analysis in Twitter," in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, 2017, doi: 10.18653/v1/S17-2088.
[26] L. Branz and P. Brockmann, "Sentiment Analysis of Twitter Data," in *Proceedings of the 12th ACM International Conference on Distributed and Event-based Systems - DEBS '18*, 2018, doi: 10.1145/3210284.3219769.
[27] A. Alrumaih, A. Al-Sabbagh, H. Kharrufa, and J. Baldwin, "Sentiment analysis of comments in social media," *Int. J. Electr. Comput. Eng.*, vol. 10, no. 6, pp. 5917–5922, 2020, doi: 10.11591/ijece.v10i6.pp5917-5922.
[28] G. A. Ruz, P. A. Henríquez, and A. Mascareño, "Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers," *Futur. Gener. Comput. Syst.*, vol. 106, pp. 92–104, 2020, doi: 10.1016/j.future.2020.01.005.
[29] D. S. Maylawati, M. N. Mudyawati, M. H. Wahisyam, and R. A. Maulana, "Comparison of Classification Algorithms for Sentiment Analysis on Movie Comments," in *Gunung Djati Conference Series*, 2021, vol. 3, pp. 68–80.
[30] C. Nanda, M. Dua, and G. Nanda, "Sentiment Analysis of Movie Reviews in Hindi Language Using Machine Learning," *Proc. 2018 IEEE Int. Conf. Commun. Signal Process. ICCSP 2018*, pp. 1069–1072, 2018, doi: 10.1109/ICCSP.2018.8524223.
[31] A. Z. Amrullah, A. Sofyan Anas, and M. A. J. Hidayat, "Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square," *J. Bumigora Inf. Technol.*, vol. 2, no. 1, pp. 40–44, 2020, doi: 10.30812/bite.v2i1.804.
[32] M. Z. Negara, G. F. Shidik, A. Z. Fanani, and E. Noersasongko, "Sentiment Analysis of Indonesian News Using Deep Learning (Case Study: TVKU Broadcast)," in *2018 International Seminar on Application for Technology of Information and Communication*, 2018, pp. 261–265.
[33] X. Fang and J. Zhan, "Sentiment analysis using product review data," *J. Big Data*, 2015, doi: 10.1186/s40537-015-0015-2.
[34] K. Marimuthu, A. Shankar, R. Ranganathan, and R. Niranchana, "Product opinion analysis using text mining and analysis," *Int. J. Smart Grid Green Commun.*, vol. 1, no. 3, p. 227, 2018, doi: 10.1504/IJSGGC.2018.091351.
[35] E. Rivo, J. de la Fuente, A. Rivo, E. Garcia-Fontan, M.-A. Canizares, and P. Gil, "Cross-Industry Standard Process for data mining is applicable to the lung cancer surgery domain, improving decision making as well as knowledge and quality management," *Clin. Transl. Oncol.*, vol. 14, no. 1, pp. 73–79, 2012.
[36] P. Chapman *et al.*, "The CRISP-DM user guide," in *4th CRISP-DM SIG Workshop in Brussels in March*, 1999, vol. 1999.
[37] B. Bachmann, "CRISP-DM siap untuk Proyek Pembelajaran Mesin," *ICHI.PRO*, 2021. https://ichi.pro/id/crisp-dm-siap-untuk-proyek-pembelajaran-mesin-46924957877610 (accessed Oct. 02, 2021).
[38] D. S. Maylawati, H. Sugilar, and R. A. Yudhiantara, "KUALITAS PERANGKAT LUNAK: MODULARITAS PUSTAKA TEXT PRE-PROCESSING," *J. Perspekt.*, 2018, doi: 10.15575/jp.v1i2.11.
[39] S. Kannan *et al.*, "Preprocessing Techniques for Text Mining," *Int. J. Comput. Sci. Commun. Networks*, 2015.
[40] M. Adriani, J. Asian, S. M. M. T. . Nazief, and H. . Williams, "Stemming Indonesian: A Confix-stripping approach," *ACM Trans. Asian Lang. Inf. Process.*, vol. 6, no. 1, pp. 1 – 33, 2007.
[41] A. Santoso and G. Ariyanto, "Implementasi Deep Learning Berbasis Keras Untuk Pengenalan Wajah," *Emit. J. Tek. Elektro*, vol. 18, no. 01, pp. 15 – 21, 2018.
[42] S. Saha, "A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way," *towardsdatascience.com*, 2018. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53.
[43] S. Sena, "Pengenalan Deep Learning Part 7 : Convolutional Neural Network (CNN)," *Medium.com*, 2017. https://medium.com/@samuelsena/pengenalan-deep-learning-part-7-convolutional-neural-network-cnn-b003b477dc94 (accessed Aug. 02, 2019).
[44] K. M. Ting, "Confusion Matrix," in *Encyclopedia of Machine Learning and Data Mining*, 2017.
[45] M. Gambhir and V. Gupta, "Recent automatic text summarization techniques: a survey," *Artif. Intell. Rev.*, 2017, doi: 10.1007/s10462-016-9475-9.
[46] R. Rautray and R. C. Balabantaray, "An evolutionary framework for multi document summarization using Cuckoo search approach: MDSCSA," *Appl. Comput. informatics*, vol. 14, no. 2, pp. 134–144, 2018.
[47] P. Verma, S. Pal, and H. Om, "A Comparative Analysis on Hindi and English Extractive Text Summarization," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 18, no. 3, pp. 1–39, 2019.
[48] J. Brownlee, "What Are Word Embeddings for Text?," *Machine Learning Mastery*, 2019. https://machinelearningmastery.com/what-are-word-embeddings/ (accessed Apr. 13, 2020).
[49] S. Banerjee, "Word2Vec — a baby step in Deep Learning but a giant leap towards Natural Language Processing," 2018. https://medium.com/explore-artificial-intelligence/word2vec-a-baby-step-in-deep-learning-but-a-giant-leap-towards-natural-language-processing-40fe4e8602ba.
[50] H. Zamani and W. B. Croft, "Relevance-based Word Embedding," in *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '17*, 2017, doi: 10.1145/3077136.3080831.
[51] M. Tezgider, B. Yildiz, and G. Aydin, "Improving Word Representation by Tuning Word2Vec Parameters with Deep Learning Model," in *2018 International Conference on Artificial Intelligence and Data Processing, IDAP 2018*, Jan. 2019, doi: 10.1109/IDAP.2018.8620919.
[52] K. K. D. Assistant, "WORD2VEC AND TF-IDF WORD EMBEDDING," vol. 5, no. 3, pp. 2348–2350, 2018.
[53] D. Liu, "A Practical Guide to ReLU," *Medium.com*, 2017. https://medium.com/@danqing/a-practical-guide-to-relu-b83ca804f1f7 (accessed May 09, 2020).