A Fuzzy Logic in Election Sentiment Analysis: Comparison Between Fuzzy Naïve Bayes and Fuzzy Sentiment using CNN

Sugiyarto[1], Joko Eliyanto[2], Nursyiva Irsalinda[3], Zhurwahayati Putri[4], Meita Fitrianawati[5]
[1,2,3,4,5]Departement of Mathematics, Ahmad Dahlan University, Yogyakarta, Indonesia
sugiyarto@math.uad.ac.id

ABSTRACT

Sentiment analysis is an analysis with an objective to identify *like, dislike, comments, opinion,* or *feedback* on certain content which will be categorized into positive, negative, or neutral. In general selection, sentiment analysis widely known to be used to predict the winner on election process. This method tries to dig the people sentiment on their governor candidates during election, whether it's positive, negative, or neutral opinion. The output of the positive sentiment is related to people acceptance towards one of the election nominee. That statement usually applied as a base reference for determining the result of the election process. In sentiment analysis, the importance of its fuzzy logics must be considered. Each of the people statement is assumed to have the level of positive, negative, or neutral percentage. The concept of fuzzy logic is developed and applied on one of this text mining method. This research is focusing on comparison analysis and fuzzy logic application in sentiment analysis method. Two method which discussed in this research are Fuzzy Naïve Bayes and Sentiment Fuzzy with convolutional neural network. This research is applied on PILKADA of Solo and Medan district case study. The data of the people opinion are acquired from *twitter* and collected on September 2020 to December 2020. The two methods which mentioned before are implemented on the acquired data and the output of these method application then compared. The conclusion of this research suggest that different approach will resulting in different output.

——————————— ◆ ———————————

A. INTRODUCTION

Nowadays, the application of social media are widely used by the people to express their ideas on recent event that occur in their society (Fitri et al., 2019). Twitter is one of the popular social media application which widely used as platform for their user to post a message known as *tweet* to the maximum limit in 280 character (Ansari et al., 2020). Indonesia is a country with one of the most active Twitter user (Alita et al., 2019). In 2020, the Regional Head Election or PILKADA is one of the topics which pique people interest in Indonesia. This PILKADA is held simultaneously in December 9th 2020. The involvement of the current Indonesia president's relatives became one of the main attraction for the people in Indonesia. The son of the current president, Gibran Rakabuming declared himself as the candidate of the head regional of Solo together with one of the nominee from PDIP, Teguh Prakosa. Meanwhile, in PILKADA Medan, Muhammad Bobby Afif Nasution, the son in law of the president of Indonesia, also declared to become one of the nominee for the regional head of Medan election. Just like the other candidates, they also uses social media to maximize their campaign during election period which accidentally held during COVID pandemic in Indonesia. Twitter become one of the main

social media which applied by the election nominee for their campaign (Bansal & Srivastava, 2018). During this campaign period, the opinion and the message from the candidates and the voter which delivered on Twitter resulting in certain unstructured data.

Sentiment analysis is one of the branches of the *text mining* process with the purpose to clarify the review data into certain classification (Lutfi et al., 2018). Furthermore, this sentiment analysis is applied to identify *like, dislike, comment, opinion,* or *feedback* on certain content which will be categorized into positive, negative, or neutral (Budiharto & Meiliana, 2018). There are a lot of study regarding the application of sentiment analysis which applied on different topics, such as, social media comments, products, politics, etc. (Purnomoputra & Wisesty, 2019). The sentiment analysis which specified on election topic known as *Election Sentiment Analysis*. For further specification from the sentiment value, the application of fuzzy logic is introduced. The sentiment analysis with further application of fuzzy logic will provide a clearer views to more accurate sentiment value. This matter will support the decision maker to make an effective decision suitable for their products or services (Rahman, 2014).

The model of the fuzzy logic is based on the fuzzy set theorem proposed by Zadeh (Zadeh, 1988). Set of fuzzy is defined as a set with the membership value ranged between "0" to "1" which describe the membership degree of certain object. If the binary logic is defined as "Yes" and "No", the data valued close to 0 is considered as "No" decision and vice versa. In sentiment analysis, the binary logic is defined as "Positive" and "Negative". In this research, "0" is associated with negative sentiment, "1" is associated with positive sentiment and the value between "0" and "1" is considered as the sentiment degree of the opinion. Moreover, the neutral sentiment is also considered, with the value of "0.5". In sentiment analysis, the application of fuzzy logic is used in three different methods. The first method is by applying fuzzy logic directly on sentiment analysis (Rahman, 2014)(Jefferson et al., 2017). The second one is by combining fuzzy method with one of traditional sentiment analysis which is the Fuzzy Naïve Bayes (Fathurochman et al., 2104). Then, the last one is by applying fuzzy logic together with machine learning Convolutional Neural Networks (Fathurochman et al., 2104)(Fathurochman et al., 2104)(Fathurochman et al., 2104).

The Fuzzy Naïve Bayes is a modification method which combine two methods, which are fuzzy logic and Naïve Bayes with the objective to provide better output. Fuzzy logic is a method to solve a problem which have certain level of ambiguity. In the system which needs complexity, fuzzy logic is difficult to apply and tends to provide long time for define the suitable rules and membership settings. Meanwhile, Naïve Bayes Classifier is a statistic classification method which have the ability to predict the probability of certain classification with the application of Bayes theorem. The theorem of Bayes is a theory which used to measure the class probability from each of the variables classification. Naïve Bayes explains simple decision, which is the relation between each variable in certain classification is fully independent with another variables. This simple assumption creates the prediction which is not quite optimum, and leads to inaccuracy of the output, especially for the domain text problems (Fathurochman et al., 2104). Tree and K-NN also state that accuracy and speed are the most supportive factors in data classification (Fathurochman et al., 2104). By combining these two methods, hopefully the application of Fuzzy Naïve Bayes will solve the ambiguity problem on forecasting the certain class membership matter. Other than the traditional method such as Naïve Bayes, another different method of sentiment analysis which applied in this research is model machine learning.

Sentiment analysis can be applied based on dictionary or machine learning (Fathurochman et al., 2104). Machine learning works based on the algorithm in the data classification which labelled manually. The data which trained by machine learning gives an impactful effect on the resulted classification model, the higher the value the better the output and vice versa. Hence, sentiment analysis based on machine learning will provides a dynamic advantage to meet the

new dataset languages or words. The approach by applying this method is better than the approach based on dictionary. One of the popular algorithm in machine learning is convolutional neural networks.

Based on the previous research, convolutional neural networks is proven to be one of the branches of deep learning which provides better level of accuracy (Fathurochman et al., 2104). However, the output of the sentiment analysis is a clear classification, which are positive, negative, and neutral (Sánchez-Rada & Iglesias, 2019). In CNN model, the classification result is acquired by selecting the highest probability value on each of the classification output. Another form of the sentiment analysis is sentiment score, where the output of this method is valued as positive, negative, or zero. However, this sentiment score is resulted by the application of sentiment analysis based on dictionary. In fact, the sentiment value on every sentences is not fully positive, negative, or neutral, there will always be ambiguity level there, and the concept of fuzzy logic is one of the good method to define that ambiguity.

This research is done based on the many research of sentiment analysis method application in a lot of problems in real life (Fitri et al., 2019), (Alita et al., 2019), (Susanti et al., 2017)(Andrianti, 2018)(Alamsyah, 2018)(Deriu & Cieliebak, 2016)(Nugroho et al., 2015), (Lutfi et al., 2018)(Budiharto & Meiliana, 2018)(Purnomoputra & Wisesty, 2019), (Rajput et al., 2016)(D, 2019)(Suryani et al., 2019)(Odim et al., 2020)(Normah, 2019). Furthermore, the development on sentiment analysis method is always be done (Rahman, 2014), (Jefferson et al., 2017)(Fathurochman et al., 2104)(Arunabala et al., 2019)(Nasichuddin et al., 2018), (Kim & Jeong, 2019)(Sutabri et al., 2018)(Suppala & Rao, 2019)(Dey et al., 2016). Based on the statement above, the fuzzy naïve Bayes sentiment analysis and fuzzy sentiment analysis with convolutional networks are chosen, applied, and compared with the objective to analyse the role of fuzzy logic in traditional sentiment method and machine learning.

B. METHODS

This research focuses on the result comparison of the sentiment analysis by Fuzzy Naïve Bayes method and Fuzzy Sentiment method on PILKADA 2020 in Solo and Medan:

1. Fuzzy Naïve Bayes

Naive Bayes Classifier

Naïve Bayes Classifier is a machine learning technique based on the theory of probability. Although this method is simple, but this method is proven to provides high level of accuracy and performance in text classification. There are two main processes in Naïve Bayes Classifier method, which are training process and testing process. The training process will result in the rules of classification which contain the probability value for each variable (conditional probability) and the targeted classification from the data train that already determined. The testing process is a testing process for each variables to determine the optimum class based on the classification rules which already acquired (Fathurochman et al., 2104).

In Naïve Bayes Classifier, every "tweet" data are defined as the combination of certain variable $(x_1, x_2, x_3, \dots, x_n)$ where x_1 is the first word, x_2 is the second word, and so on. Then, there are V value which defined as the set of classification. After the pre-processing step already done, the classification process in Naïve Bayes Classifier is applied to clarify the data before further step is initiated. There are 4 processes which applied on classification process, such as:

a. Establishment of trained data variable

This process is an initial step in Naïve Bayes Classifier classification process, where the mentioned variable is described as tweet (comments) from Twitter which will be classified into sentiment class; positive, negative, or neutral by observing the emerged frequency on every words in the variable.

b. Measurement of the trained data probability value

After the trained data variable are established, its probability is measured with equation as follows:

$$P(v_j) = \frac{n(v_j)}{n(S)} \tag{1}$$

c. Determination of class probability

In this process, the probability in each classes can be measured by applying this following equation:

$$P(x_I|v_j) = \frac{n_i+1}{n+f_k} \tag{2}$$

$n(v_j)$ is the number of classes in category of j, where j define the row order of the data. $n(S)$ is the number of classes on trained data.

d. Measurement of V_{MAP}

In classification process, this method will provide a class with the highest probability value (V_{MAP}) from the tested data in each class for every variable combination $(x_1, x_2, x_3, \ldots, x_n)$. The value of maximum probability (V_{MAP}) is determined by using the following equation:

$$(V_{MAP}) = \frac{\arg\,max}{v_j} P(v_j) \prod_{i=1}^{n} f(x_i|v_j) \tag{3}$$

Fuzzy Association

If X is a set of object which generally notated as x, then the set of *fuzzy* $\tilde{A}$ in X is the row combination of:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x))|x \epsilon X\} \tag{4}$$

$\mu_{\tilde{A}}(x)$ is defined as the membership function of x in set of M between 0 to 1 (Honda & Ohsato, 1986).

Fuzzy Membership Function

The membership function is defined as a curve which shows the pattern of every input data coordinates into its membership value (membership degree) which ranged from 0 to 1 interval (Honda & Ohsato, 1986). One of the function that can be applied is the triangle curve.

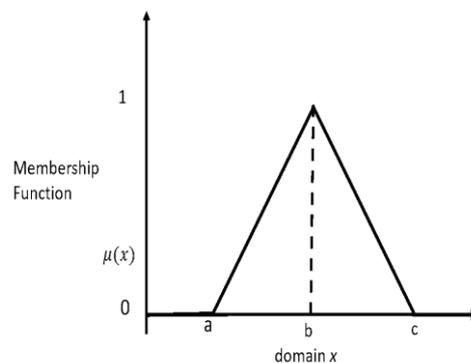The triangle curve is a combination of two linear lines as described in Figure 1.

Figure 1. Triangle Curve

Membership degree:

$$\mu(x) = \begin{cases} 0; & x \le a \text{ or } x \ge c \\ \frac{(x-a)}{(b-a)}; & a < x \le b \\ \frac{(c-x)}{(c-b)}; & b < x < c \end{cases} \tag{5}$$

Confusion Matrix

Matrix confusion can be formed by using this following equation:

$$Confusion\ Matrix = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix} \tag{6}$$

Based on that equation, accuracy level by applying 2 or more *n* class is (Manliguez, 2016):

$$Accuracy = \frac{\sum_{j=1}^{j=n} x_{i,j}}{number\ of\ testing\ data} \tag{4}$$

2. Fuzzy Sentiment with Convolution Neural Networks

Fuzzy sentiment method by convolutional neural network tries to create sentiment value in the form fuzzy, which is between 0 and 1. This method is included as one of the part in machine learning process. The initial steps to apply in this method are data cleaning, tokenization, encoding, and padding. This process is described in figure 2. Data cleaning is a process where the unwanted data from the words or characters which being stated in the opinion are cleaned. This purpose of this process is to diminish the data which will affect the output of sentiment analysis process. The next step is tokenization, this step means to break the sentences into single words. Then, this single words will be encoded into numbers to be structured into sentence vector. The length difference in each sentences then will be uniformed by applying padding method. After the padding process is finalized, the measurement of the sentiment is applied by using Convolutional Neural Network method.
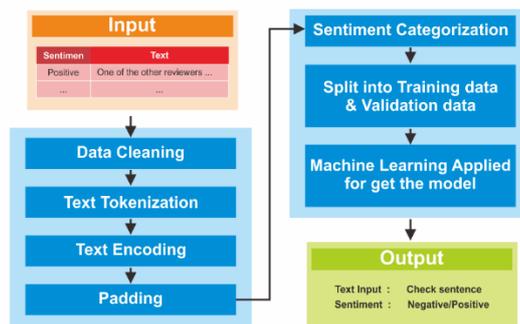

Figure 2. Machine Learning Approah for Sentiment Analysis

The structure of convolutional neural networks is consisted into 3 main layers, which are input layer, feature extraction layer, and classification layer (Nasichuddin et al., 2018). This structure is described in figure 3. For text classification, embedding process is applied in input layer. The words insertion is a process to interpret the words as a bulk vector. This method is applied to identify the adjacency between words through vector representation. This process will be described thoroughly in figure 4.
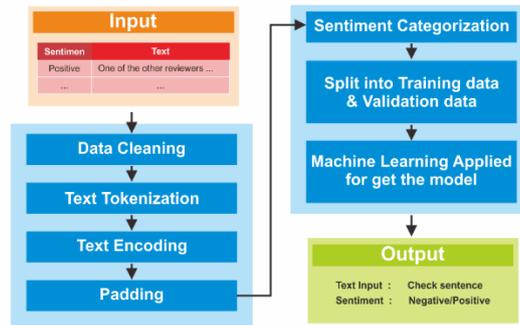
Figure 3. Convolutional Neural Networks

In the next layer, convolution process will be applied. The term of convolution refers to mathematics combination from 2 function to form the next third function. The objective of this process is for feature extraction so that the only features which will be considered in neural networks model are the important one.
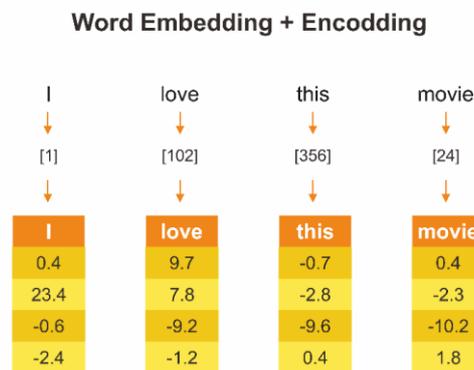

Figure 4. Word Embedding

To prevent overfitting, dropout process is applied. Dropout is a technique where the unnecessary neurons which are randomly chosen will be ignored during training (Melucci, 2016). This means that their contribution towards downstream neuron activation is eliminated temporarily in forward pass process and every weight updating process won't be affected by neuron which applied in backward pass. Dropout structure will be shown in figure 5. Maxpooling process then will be applied to reduce the computational weight of this algorithm. For each classification, the suitable activation function is applied. For binary classification, sigmoid activation function will be used while softmax activation function is applied for multiclass classification. The result of this modelling process is probability value on each class classification. Class featured with the highest value of probability will be considered as the referred object.
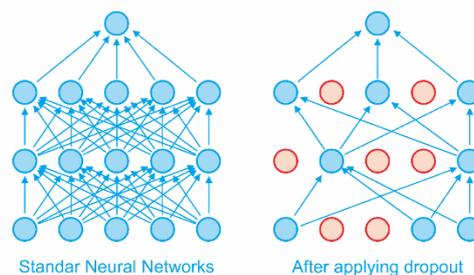

Figure 5. Dropout

Then, these result will be modified into set of fuzzy. The modification of model result into set of fuzzy will be mentioned on figure 6(a) for binary class and figure 6(b) for multiclass. The output of this method are vector of sentences sentiment value on dataset. The similarity in this vector will be measured by applying Euclid distance concept with dataset label sentiment vector. The Euclid distance between two vectors, $\vec{p} = (p_1, p_2, \cdots, p_n)$ and $\vec{q} = (q_1, q_2, \cdots, q_n)$ inside Euclidean space is defined as follows (Strang, 2016):

$$d(\vec{p}, \vec{q}) = d(\vec{q}, \vec{p})$$

$$= \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \tag{8}$$

The distance between fuzzy sentiment vector and standard convolutional neural networks sentiment vector will be compared. The distance result which is lower means that this vector is close to the dataset label sentiment vector. This concludes that the applied sentiment value is better, or close to the real output.
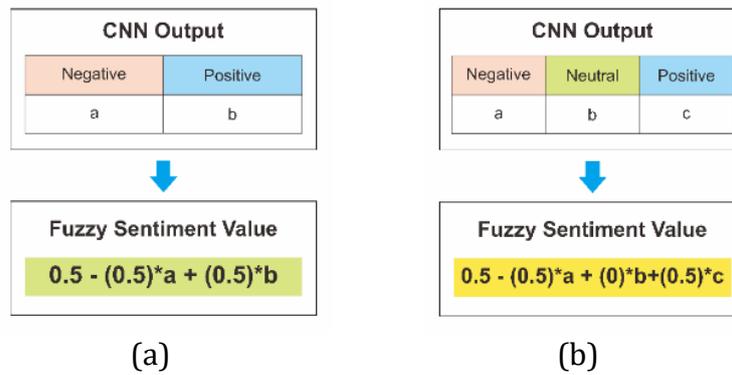

(a) (b)

Figure 6. Create *Fuzzy* Sentiment Value

3. Dataset

Dataset is defined as the tweets which are taken from twitter starting from 26 September 2020 to 6 December 2020. In PILKADA Solo, the two candidates who compete are Gibran Rakabuming – Teguh Prakosa as candidate one, and Bagyo Wahyono – FX Supardjo as candidate two (KPUD Kota Surakarta, 2020). The applied keywords for this parameter are Gibran-Teguh, and Bagyo-Supardjo. In addition to that, in PILKADA Medan(KPUD Kota Medan, 2020), the competing candidates are Akhyar Nasution – Salman Al-Farisi as candidate one, and Muhammad Bobby Afif Nasution – Aulia Rahman as candidate two. The applied keywords for this parameter are Akhyar-Salman, and Bobby-Aulia. The number of tweets which collected are described in Table 1.

Table 1. Crwaling Data Result

Keyword	Number of Tweets
gibran-teguh	5.259
bagyo-supardjo	119
akhyar-salman	7.340
boby-aulia	1.962

C. RESULT AND DISCUSSION

Fuzzy Naïve Bayes method resulting on the percentage value of sentiment for each candidate. This percentage is illustrated in figure 7. Based on the output data in figure 7, candidate Gibran-Teguh receive 27.915% positive sentiment, 55.008% negative sentiment, and 17.077% neutral sentiment with level of accuracy valued 66.9%. Meanwhile, for candidate Bagyo-Supardjo, resulting in 33.333%, 54.167%, and 12.500% for positive sentiment, negative sentiment, and neutral sentiment respectively with accuracy percentage of 54.16%. Furthermore, for PILKADA Medan, candidate Akhyar-Salman received 21.577% positive sentiment, 63.900% negative sentiment, and 14.523% neutral sentiment in 63.48% accuracy. For candidate 2, Bobby-Aulia received 26.718%, 50,725%, and 17.557% for positive, negative, and neutral sentiment respectively in 67.17% accuracy. These data mentioned before shows that the highest sentiment percentages received for four nominees is the negative ones.
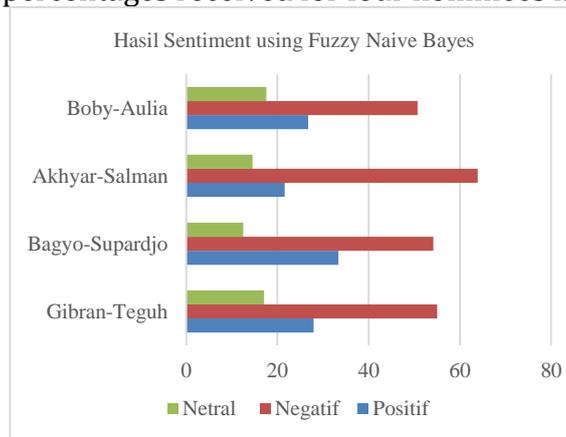

Figure 7. Result of Sentiment using *Fuzzy* Naïve Bayes

Next, the average result for the sentiment fuzzy with convolutional neural network is illustrated on table 2. While for the trend of sentiment value for each week is shown on figure 8. Based on these data, the sentiment value for candidates in PILKADA Solo tends to be positive with the acquired value higher than 0.5. Meanwhile, the sentiment value for candidates in PILKADA Medan tends to be negative with the acquired value lower than 0.5. Nevertheless, the last acquired result actually linear with the Fuzzy Naïve Bayes method. Besides that, the transformation of the sentiment for each week in campaign period is also fluctuates.

Table 2. Average Result of *Fuzzy* Sentiment Score CNN

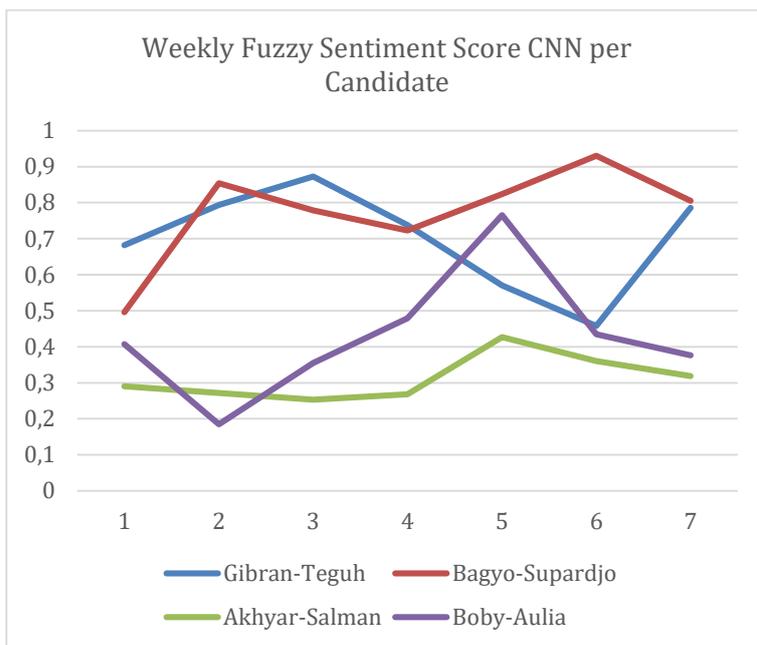| Pair of Candidates	Average Result og *Fuzzy* Sentiment Analysis
Gibran-Teguh	5.259
Bagyo-Supardjo	119
Boby-Aulia	7.340
Akhyar-Salman	1.962

Figure 8. Weekly *Fuzzy* Sentiment Score CNN per Candidate

The result of the PILKADA are already announced, candidate Gibran-Teguh won the election in Solo with 225.451 % vote defeating candidate Bagyo-Supardjo with 35.055% vote [41]. In Medan, candidate Akhyar-Salman receive 324.480 votes, losing by a neck to its opponent, candidate Bobby-Aulia who acquired 393.327 votes. The positive sentiment resulted from Fuzzy Naïve Bayes and Fuzzy Sentiment CNN then compared with the real count result from KPU (General Election Committee) which illustrated on figure 9.
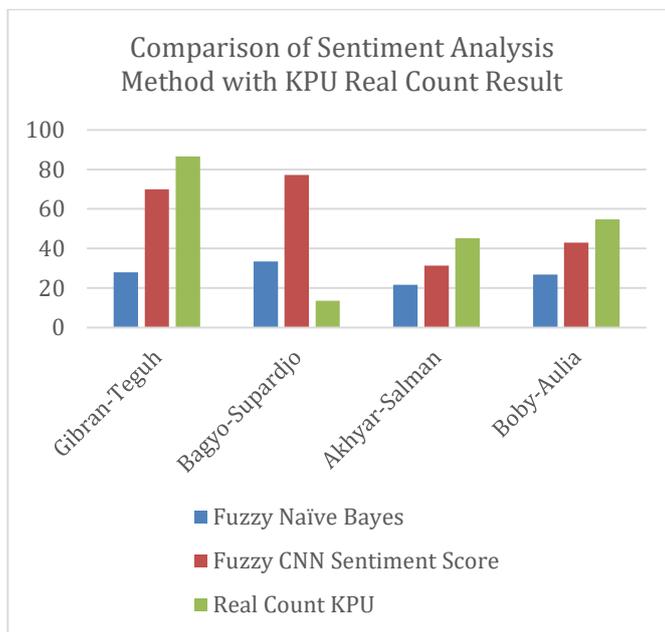

Figure 9. Comparison of Sentiment Analysis Method with KPU Real Count Result

On this comparison, for the candidate in PILKADA Medan, the one who receive higher positive sentiment percentage in Fuzzy Sentiment CNN and Fuzzy Naïve Bayes method successfully win the election. Meanwhile, the opposite comparison achieved for PILKADA Solo, the candidate who receive lower positive sentiment percentage in Fuzzy Sentiment CNN and Fuzzy Naïve Bayes method, in fact acquiring higher number of votes and win in a huge margin. This phenomenon occur because the twitter data for the losing candidate which are collected for this research are too low, only 119 tweets. This conclusion concludes by considering the

concept of Machine Learning where the sentiment analysis process will works better in bigger dataset.

D. CONCLUSION AND SUGGESTIONS

This research concludes that the two applied method are well implemented. The concept of fuzzy logic have the ability to solve the ambiguity problem in sentiment analysis. By applying huge number of data, Fuzzy Naïve Bayes and Fuzzy Sentiment with Convolutional Neural Network could generate the sentiment value which represent the real count result in PILKADA. In the future, the application of fuzzy logic should be perfected in sentiment analysis method. This statement is important considering that ambiguity is the main and general problem to determine the sentiment from certain statement.

REFERENCES

Alamsyah, D. M. G. A. (2018). Employee Sentiment Analysis Using Naive Bayes Classifier. *Int. J. Sci. Res*, *7*(5), 1676–1679.

Alita, D., Priyanta, S., & Rokhman, N. (2019). Analysis of Emoticon and Sarcasm Effect on Sentiment Analysis of Indonesian Language on Twitter. *J. Inf. Syst. Eng. Bus. Intell.*, *5*, 100.

Andrianti, S. D. (2018). *Sentiment Analysis On Twitter Account Using Naive Bayes Classifier Algorithm Case Study : Indonesia Healthcare and Social Security Agency (BPJS Kesehatan)*. *2*(2), 23–28.

Ansari, M. Z., Aziz, M. B., Siddiqui, M. O., Mehra, H., & Singh, K. P. (2020). Analysis of Political Sentiment Orientations on Twitter. *Procedia Comput. Sci*, *167*, 1821–1828.

Arunabala, D. C., Jwalitha, P., & Nuthalapati, S. (2019). Text Sentiment Analysis Based on Cnns and Svm. *Int. J. Res. -GRANTHAALAYAH*, *7*(6), 77–83.

Bansal, B., & Srivastava, S. (2018). On predicting elections with hybrid topic based sentiment analysis of tweets. *Procedia Comput. Sci*, *135*, 346–353.

Budiharto, W., & Meiliana, M. (2018). Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis. *J. Big Data*, *5*(1), 1–10.

D, C. V. (2019). Hybrid approach: naive bayes and sentiment VADER for analyzing sentiment of mobile unboxing video comments. *Int. J. Electr. Comput. Eng*, *9*, 5.

Deriu, J., & Cieliebak, M. (2016). Sentiment analysis using convolutional neural networks with multi-task training and distant supervision on Italian tweets. *CEUR Workshop Proc*, *1749*.

Dey, L., Chakraborty, S., Biswas, A., Bose, B., & Tiwari, S. (2016). Sentiment Analysis of Review Datasets Using Naïve Bayes' and K-NN Classifier. *Int. J. Inf. Eng. Electron. Bus*, *8*(4), 54–62.

Fathurochman, D., Witanti, W., & Yuniarti, R. (2104). Perancangan Game Turn Based Strategy Menggunakan Logika Fuzzy dan Naive Bayes Classifier. *Semin. Nas. Inform*, 44–51.

Fitri, V. ., Andreswari, R., & Hasibun, M. A. (2019). Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm. *Procedia Comput. Sci*, *161*, 765–772.

HONDA, N., & OHSATO, A. (1986). *Fuzzy Set Theory and Its Applications*. *13*(2).

Jefferson, C., Liu, H., & Cocea, M. (2017). Fuzzy approach for sentiment analysis. *IEEE Int. Conf. Fuzzy Syst*.

Kim, H., & Jeong, Y. S. (2019). Sentiment classification using Convolutional Neural Networks. *Appl. Sci*, *9*(11), 1–14.

KPUD Kota Medan. (2020). *Salinan Sk N0. 1672 Tentang Penetapan Rekapitulasi Hasil*

Penghitungan Perolehan Suara Dan Hasil Penetapan Walikota Dan Wakil Walikota Medan Tahun. KPUD Kota Medan.

KPUD Kota Surakarta. (2020). *SK No.212 Rekap Hasil Perolehan Suara*. KPUD Kota Surakarta.

Lutfi, A. A., Permanasari, A. E., & Fauziati, S. (2018). Corrigendum: Sentiment Analysis in the Sales Review of Indonesian Marketplace by Utilizing Support Vector Machine. *J. Inf. Syst. Eng. Bus. Intell*, *4*(2), 169.

Manliguez, C. (2016). *Generalized Confusion Matrix for Multiple Classes Generalized Confusion Matrix for Multiple Classes The total numbers of false negatif (TFN), false positive (TFP), and true negatif (TTN) for each class i will be calculated based on the Generalized*. 5–7.

Melucci, M. (2016). Relevance Feedback Algorithms Inspired by Quantum Detection. *IEEE Trans. Knowl. Data Eng*, *28*(4), 1022–1034.

Nasichuddin, M. A., Adji, T. B., & Widyawan, W. (2018). Performance Improvement Using CNN for Sentiment Analysis. *IJITEE (International J. Inf. Technol. Electr. Eng*, *2*(1).

Normah, N. (2019). Naïve Bayes Algorithm For Sentiment Analysis Windows Phone Store Application Reviews. *SinkrOn*, *3*(2), 13.

Nugroho, D. G., Chrisnanto, Y. H., & Wahana, A. (2015). *Analisis Sentimen Pada Jasa Ojek Online*. 156–161.

Odim, M. O., Ogunde, A. O., Oguntunde, B. O., & Phillips, S. A. (2020). Exploring the performance characteristics of the naïve bayes classifier in the sentiment analysis of an Airline's social media data. *Adv. Sci. Technol. Eng. Syst*, *5*(4), 266–272.

Purnomoputra, R. B., & Wisesty, U. N. (2019). Sentiment Analysis of Movie Reviews using Naïve Bayes Method with Gini Index Feature Selection. *Open Access J Data Sci Appl*, *2*(2), 85–094.

Rahman, T. (2014). *S a u f l. 4*(1), 33–48.

Rajput, Q., Haider, S., & Ghani, S. (2016). Lexicon-Based Sentiment Analysis of Teachers' Evaluation. *Appl. Comput. Intell. Soft Comput*, 1–12.

Sánchez-Rada, J. F., & Iglesias, C. A. (2019). Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison. *Inf. Fusion*, *52*, 344–356.

Strang, G. (2016). Introduction to Linear Algebra, Fifth Edition. In *Wellesley-Cambridge Press*.

Suppala, K., & Rao, N. (2019). Sentiment analysis using naïve bayes classifier. *Int. J. Innov. Technol. Explor. Eng*, *8*(8), 264–269.

Suryani, P. S. M., Linawati, L., & Saputra, K. O. (2019). Penggunaan Metode Naïve Bayes Classifier pada Analisis Sentimen Facebook Berbahasa Indonesia. *Maj. Ilm. Teknol. Elektro*, *18*(1), 145.

Susanti, A. R., Djatna, T., & Kusuma, W. A. (2017). Twitter's sentiment analysis on GSM services using Multinomial Naïve Bayes. *Telkomnika (Telecommunication Comput. Electron. Control*, *15*(3), 1354–1361.

Sutabri, T., Suryatno, A., Setiadi, D., & Negara, E. S. (2018). Improving naïve bayes in sentiment analysis for hotel industry in Indonesia. *Proc. 3rd Int. Conf. Informatics Comput. ICIC 2018*, 1–6.

Zadeh, L. A. (1988). Fuzzy Logic. In *Computer (Long. Beach. Calif)*.