Implementation of the Convolutional Neural Network (CNN) Method for Sentiment Analysis in Teaching and Learning Process Evaluation (PBM)

I Wayan Suasnawa[1], I Gusti Ngurah Bagus Catur Bawa[2],
Anak Agung Ngurah Gde Sapteka[3], Ni Gusti Ayu Putu Harry Saptarini[4],
Ida Bagus Putra Manuaba[5], and I Komang Wiratama[6]

[1,2,4,5,6] Information Technology Department, Politeknik Negeri Bali, Bali, Indonesia
[3] Electrical Engineering Department, Politeknik Negeri Bali, Bali, Indonesia
suasnawa@pnb.ac.id

Abstract. This research aims to apply the Convolutional Neural Network (CNN) method in analyzing sentiment related to the teaching and learning process (PBM). PBM is a critical aspect in the world of education, and evaluation of this process can provide valuable insight for decision making. The CNN method is used to process text data containing PBM evaluations from student surveys. The steps in this research include data collection, data preprocessing, text representation, CNN architecture, training model, and evaluation model. It is hoped that the results of this research will provide insight into sentiment regarding PBM so that educational institutions can take appropriate action to improve the quality of the teaching and learning process. The CNN algorithm obtained an overall accuracy value of 72% with average negative, neutral, and positive precision values of 93%, 53%, and 83% respectively. This shows the ability to analyze sentiment is still lacking in neutral sentiment analysis.

Keywords: Convolutional Neural Network (CNN), Learning Process, PBM

1 Introduction

The teaching and learning process (PBM), also known as the learning process, is a combination of two concepts, namely learning carried out by students and teaching carried out by lecturers. Implementing PBM is an important factor that must be carried out by universities. In order for PBM to run smoothly, it is necessary that universities should pay special attention to this matter (Wulandari et al., 2023).

Of course, there are some obstacles faced by teachers and students during PBM (Sugiarta et al., 2023). There are many comments both good and bad regarding the implementation of this PBM. For this reason, it is necessary to know the extent of student sentiment regarding the implementation of PBM. Student sentiment can be grouped into two, those who feel that the implementation of PBM has been carried out well and some feel that the learning atmosphere is less supportive, the learning foundation is not strong enough, the learning environment is less conducive, the

A. A. N. G. Sapteka et al. (eds.), *Proceedings of the International Conference on Sustainable Green Tourism Applied Science - Engineering Applied Science 2024 (ICoSTAS-EAS 2024)*, Advances in Engineering Research 249,
https://doi.org/10.2991/978-94-6463-587-4_9

teaching design and delivery of lesson material. To find out student sentiment, we can take and use survey data on students. Surveys are a means of conveying opinions on campus which can be used as reference material for student sentiment towards PBM (Sunardi et al., 2018). With methods in the computer world known as machine learning, results can be obtained in the form of negative comments, neutral comments, and positive comments. In this research, the Convolutional Neural Network (CNN) method will be used to obtain sentiment analysis regarding the implementation of PBM at the Bali State Polytechnic. Although sentiment analysis has been applied extensively in many fields, traditional machine learning techniques like K-Nearest Neighbors and Naive Bayes have been the mainstay of earlier research on educational assessment (Sugiarta et al., 2023; Sunardi et al., 2018). In the teaching and learning process (PBM), this study presents a novel method for sentiment analysis of student feedback using Convolutional Neural Networks (CNN). Automatic feature extraction is made possible by CNN's deep learning capabilities, which may lead to more precise insights regarding student sentiment. As far as we are aware, this is the first study at Bali State Polytechnic to use CNN for PBM evaluation, which represents a substantial development in the field of sentiment analysis in education.

2 Methodology

2.1 Data Collection

The data collection carried out was in the form of collecting data related to the data needed as initial data, namely student comments regarding PBM that they felt while attending lectures at Bali State Polytechnic. Data was collected during a student survey.

2.2 Data Preprocessing

Data Preprocessing is a stage that must be carried out before proceeding to the data learning or training stage. This stage contains the labeling process and also data cleaning. The labeling stage is giving sentiment values to sentences or comments in the form of sentiments with positive, neutral, and negative polarities. This process is carried out manually which will later be learned by the machine in the form of sentiment analysis learning. Because the data collected has missing or non-standard words, the next stage is data cleaning, namely the process of preparing words or sentences so they can be processed in data learning. This stage is very useful to make it easier or maximize the machine in classifying existing sentences into sentiment polarity. Several stages in data cleaning are case folding, tokenizing, pad sequence, stemming, removing punctuation, and filtering/stopword removal. The data cleaning process uses the help of supporting libraries in the Python language such as NLTK, TensorFlow, Sastrawi, Re, and several other supporting libraries.

2.3 Data Extraction

At this stage, data extraction is used so that machines can read the comment data by changing the comment data in the form of text into vectors with fixed dimensions, making it easier for machines to process the data. This study includes a Convolutional Neural Network (CNN) model based on deep learning, in contrast to prior sentiment analysis approaches that mainly used classical machine learning algorithms like TF-IDF for text representation (Najiyah et al., 2021). With its convolutional and pooling layers, CNN can automatically extract features, negating the need for human feature engineering. Because of this strategy, the model's capacity to capture more complicated sentiment patterns in student feedback is improved, which is a considerable advance over standard methods. Using the CNN model, which has a layered architecture consisting of embedding, convolutional, and pooling layers, is a novel way to evaluate PBM and offers a more sophisticated sentiment analysis procedure. In this research, the data extraction was handled by the layer of CNN, such as the embedding layer, convolutional layer, pooling layer, and dense layer that was built using the TensorFlow library.

2.4 Sentiment Analysis

At this stage, the data can be processed to look for information in the form of sentiment with several types of polarity contained in the data set. The algorithm used in this sentiment analysis, namely the Convolutional Neural Network, will be tested in terms of its ability to analyze sentiment. Convolutional Neural Network is an algorithm that is part of deep learning, the way it works is to carry out data training on large data sets using parameters that take a 1-dimensional form as input. Then combine the input with a filter to get the desired output.

In general, CNN is divided into 2 layers, the extraction layer and the classification layer. The extraction layer is at the beginning of the architecture which is composed of a collection of layers where this layer consists of neurons that are connected to the previous section. Meanwhile, other classification layers do not have a section that separates neurons (Hermanto et al., 2021).

2.5 Data Splitting

At this stage, the data was split into two parts. Features and targets are the two components of the data. after that, a machine learning technique called "train-test split" divides a dataset into two parts: a train set, which is used to train the model, and a test set, which is used to assess the performance of the trained model. By ensuring that the trained model can generalize well on data that it has never seen before, this strategy helps to improve the predictability of the model's performance on new data.

The model is trained using the trainset, in this study, CNN was the model used. The training set taught the model how to identify patterns and correlations to make a prediction. The trained model, which was trained with the training set, is assessed using the test set. The training set is used to determine the model's accuracy and predictive power concerning the testing set's target.

In this research, the data is split into 80% for the training set and 20% for the testing set. The data-splitting process was carried out using the Python library Scikit-Learn (Medar et al., 2019).

2.6 Model Evaluation

Currently, a few model assessment techniques are used to assess the model. When a machine learning model is evaluated, its performance is compared on a different dataset than the one that was used for training. This process is known as model evaluation. Ensuring the model's proper generalization to new, untested data is crucial. The model evaluation techniques employed in this study are recall, precision, F1-score, and accuracy (Raschk et al., 2018).

Accuracy. Accuracy is used to prove the performance of the algorithm used using the following Equation 1 (Maulidah et al., 2020).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \tag{1}$$

Precision. Precision can be defined as the ratio of true positive predictions (TP) to the total number of positive predictions made. To calculate the precision of the mode use the Equation 2 (Maulidah et al., 2020).

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{2}$$

Recall. The ratio of true positive predictions (TP) to all true positive data is known as recall. To calculate the recall of the model use the Equation 3 (Maulidah et al., 2020).

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

F1-Score. The F1-Score compares precision and recall using a weighted average. To calculate the F1-Score of the model, use the Equation 4 (Maulidah et al., 2020).

$$\mathrm{F1 - Score} = 2 \times \frac{Precision \times Recall}{(Precision + Recall)} \tag{4}$$

3 Result and Discussion

3.1 Result

In the machine learning training process, results are obtained in the form of precision and accuracy. This parameter is used as a benchmark for the system's reliability to provide more accurate results. Testing was carried out to determine the precision and accuracy of the Convolutional Neural Network (CNN) method.

The overall accuracy that has been obtained is 0.72. The results of the procession that has been obtained from the system show that the Convolutional Neural Network (CNN) method has precision measurement results, amounting to 0.93 for negative sentiment, 0.53 for neutral sentiment, and 0.83 for positive sentiment as shown in Table 1. Based on Table 1, the recall of the model that has been obtained for each class negative, neutral, and positive is 0.95, 0.80, and 0.56. The f1-score of the model that has been obtained for each class negative, neutral, and positive is 0.94, 0.64, and 0.67.

Table 1. Test result

	Precision	Recall	f1-score
Negative	0.93	0.95	0.94
Neutral	0.53	0.80	0.64
Positive	0.83	0.56	0.67

3.2 Discussion

From the data obtained, the test results show very good precision for negative sentiment, namely 93%, and positive sentiment 83%, while neutral sentiment is still around 53%. For the recall, the negative sentiment has the best score, which is 95%. And for the f1 score, the negative sentiment has the best score, which is 94%. The overall accuracy of the model is 72%. The study’s findings highlight the special benefits of utilizing CNN for sentiment analysis in assessments of educational programs, especially in light of the technology's 93% and 83% precision scores, respectively, in classifying positive and negative sentiments. The accuracy of 53% for the neutral sentiment, however, indicates that more work needs to be done, indicating a topic for future study. These results contrast with those of earlier studies Sugiarta et al. (2023) and Sunardi et al. (2018) that used conventional techniques like Naive Bayes, which frequently had difficulty with the same classification tasks. This work adds a new approach to educational data analysis by enhancing sentiment classification in PBM and provides insights into possible areas for model improvement, especially in neutral sentiment analysis.

4 Conclusion

Based on the tests carried out, the CNN algorithm with the Global Max Pooling layer obtained an overall accuracy value of 72% with average negative, neutral and positive precision values of 93%, 53%, and 83% respectively, which shows the ability to analyze sentiment regarding the implementation of teaching and learning process (PBM) at the Bali State Polytechnic is still lacking in neutral sentiment analysis. In comparison to more established machine learning models like Naive Bayes and K-Nearest Neighbors, this study offers a novel application of Convolutional Neural Networks (CNN) for sentiment analysis in the context of educational evaluations (PBM) (Sugiarta et al., 2023; Sunardi et al., 2018; Rachman, 2021). While neutral sentiments are still difficult to identify, CNN's capacity to automatically extract characteristics from text input has shown useful in assessing both positive and negative sentiments. This work opens the door for further investigations that may further improve the use of CNN in sentiment analysis in educational contexts by showing how accurate it can be.

References

Hermanto, D. T., Setyanto, A., & Luthfi, E. T. (2021). Algoritma LSTM-CNN untuk binary klasifikasi dengan word2vec pada media online. *Creative Information Technology Journal, 8*(1), 64–77.

Maulidah, M., Gata, W., Aulianita, R., & Agustyaningrum, C. I. (2020). Algoritma klasifikasi decision tree untuk rekomendasi buku berdasarkan kategori buku. *Jurnal Ilmiah Ekonomi dan Bisnis, 13*(2), 89-96.

Medar, R., & Kaur, P. (2019). Impact of training and testing data splits on accuracy of time series forecasting in machine learning. *Journal of Machine Learning Research, 20*, 1-18.

Najiyah, I., & Haryanti, I. (2021). Sentimen analisis COVID-19 dengan metode probabilistic neural network dan TF-IDF. *Jurnal Responsif : Riset Sains Dan Informatika, 3*(1), 100–111

Prechelt, L. (1997). *PROBEN1 - A set of neural network benchmark problems and benchmarking rules*. University of Karlsruhe, Germany.

Rachman, F. P. (2021). Perbandingan model deep learning untuk klasifikasi sentiment analysis dengan teknik natural languange processing. *Jurnal Teknologi Dan Manajemen Informatika*, *7*(2). https://doi.org/10.26905/jtmi.v7i2.6506

Raschka, S. (2018). *Model evaluation, model selection, and algorithm selection in machine learning*. arXiv preprint arXiv:1811.12808.

Sugiarta, I. K. A, Dewi, P.A.C., & Utami, N. W. (2023). Analisa sentimen mahasiswa terhadap layanan Stmik Primakara menggunakan algoritma Naive Bayes dan K-Nearest Neighbor. *Jurnal Informatika Teknologi Dan Sains (Jinteks), 5*(3), 364-372.

Sunardi, Fadlil, A., & Suprianto. (2018). Analisis sentimen menggunakan metode Naãve Bayes Classifier pada angket mahasiswa. *SAINTEKBU, 10*(2), 1–9.

Wulandari, A., Salsabila, A., Cahyani, K., Nurazizah, T., & Ulfiah, Z. (2023). Pentingnya media pembelajaran dalam proses belajar mengajar. *Journal on Education, 5*(2), 3928-3936.