
Sentiment Analysis of Hotel User Review Using RNN Algorithm

Theresia Arwila Utami¹, I Wayan Ordiyasa², Hamzah³

Abstract

Sentiment analysis in user review is a growing research area at the current time. Usually, the website becomes a source of data in knowing the quality of the hotel services, and the provider can utilize the review for monitoring and evaluation. However, determining the positive or negative sentiment of a user review in unstructured textual data takes a long time. As a result, we present a model to classify positive or negative sentiment in user reviews in this article. This study suggests the RNN method in building an effective model to classify user sentiment. Based on the experiment, our model can produce accurate results in organizing hotel reviews. Furthermore, the proposed method achieved a higher evaluation metrics score with an f1-score of 91.0%.

Keywords:

Sentiment Analysis, Hotel, Classification, Deep Learning, Recurrent Neural Network

This is an open-access article under the [CC BY-SA](#) license



1. Introduction

A hotel is one of the essential tourism elements because many tourists or people traveling need a place to stay and consider in terms of facilities. A hotel has several aspects, namely food, service, comfort, and cleanliness [1]. Understanding consumers via ratings and reviews is one method to enhance the hotel business. Customer feedback on hotel services and goods plays a role in determining how to improve hotel services [2].

People who travel typically want the best hotel suggestions that meet their demands. High hotel rental costs do not ensure excellent service that meets their needs. They want the best hotel service at a fair price, but several classes are available [3]. Social media websites are a valuable source of information regarding user evaluations in various industries [4]. They're all searching for customer feedback online. Consumers give their comments based on the services they get; select hotels post their thoughts on their websites, and those views may be compared and analyzed with those of other hotels [5].

The Internet is helpful in hotel promotion in the field of hospitality tourism. Tourists often write reviews on the internet to share their hotel experiences. As a result, numerous hotel reviews can be obtained online. The effect on hotel owners is that they may use online reviews to enhance and assess their establishments. Travelers may find it difficult to understand all of the internet evaluations they read, which contain both good and negative emotions, due to the large quantity of them [6].

Various communities have proposed sentiment analysis by using different approaches, including real-world scale data [7], online reviews based on the several factors that are employed [8], or using distributed word representations for aspect-based sentiment analysis in hotel reviews [9]. Automatic labeling of text data is challenging when unstructured heaps of textual data and a significant volume of data. People frequently express their thoughts in complex and sometimes difficult-to-understand ways. Thus, it

requires a technique to extract essential features in the test dataset, such as pre-processing data [10].

In the sentiment analysis problem, online reviews effectively reveal various opinions and attitudes. The customer ratings are widely adopted to construct the benchmark dataset to deal with the case [11]. Analyzing user sentiment can identify elements that help hotel providers understand user behavior more, encourage more consumers to use their services, and anticipate future behavior [12]. The primary goal of sentiment analysis is to determine if an output is good or negative, as well as to look for differences in the text or user-specific data in a dataset [13].

The current paper explores several techniques for obtaining information using learning algorithms. The large number of reviews submitted makes the hotel need more effort and time in manual polarity labeling. A study proposed an RNN algorithm to classify user sentiment [14], using two classifications (positive and negative). Experiments were carried out using training and testing datasets. The results show that the model presents outstanding results and reaches about 91.9% [26]. The deep learning method gives good results in various implementations in the field, especially in Sentiment Analysis.

The learning method can provide high accuracy in several studies [15]. One study, for example, proposed the Attention-Based Bidirectional Deep CNN-RNN Model (ABCDM). ABCDM will extract the past and future context by analyzing the temporal information flow in both directions. The effectiveness of ABCDM is determined by sentiment polarity detection, which is the most common and important task in sentiment analysis. According to sentiment analysis, ABCDM achieved current results in both the long review and the short polarity classification [16].

The goal of sentiment classification is to determine whether a statement is positive or negative. Deep learning has proven to be a promising solution for text mining issues in the current years, including sentiment analysis categorization [17]. Based on the paper review, sentiment classification models using deep learning techniques outperform traditional machine learning models [18].

In this paper, we construct a sentiment classification model using the hotel dataset with positive and negative labels. This study has several contributions to sentiment classification using the RNN algorithm as follows:

1. We present a novel categorization model for hotel customer reviews based on RNN. In the training and testing process, we compute accuracy and loss to assess model performance.
2. We utilize deep learning to create a model of hotel user sentiment analysis to evaluate hotel quality based on positive and negative evaluations from hotel users. We also provide a graph and a confusion matrix to illustrate the model's effectiveness.
3. Based on the experiment results, the classification model can help the hotel to catch user opinions faster than the conventional model. So, the proposed model can classify hotel review sentiments accurately.

Organization: The rest of this paper is prepared as follows: Part II provides insights into related work. Part III explains the problem definition of this study. Part IV describes the experimental arrangements consisting of feature learning techniques, datasets, and data preprocessing, and Part V presents the results and detailed analysis of this study. Part VI provides a sentiment analysis conclusion for hotel users using the Recurrent Neural Network algorithm.

2. Related Works

Several studies presented several approaches for sentiment analysis, including rule-based, statistics, and learning techniques. Conventional systems proposed aspect-based multi-labeling models and evaluated them on filtered data using ensemble SVM models based on SVMs [19]. The comparative study Aspect-Based Sentimental Analysis of Hotel Review proposes a cutting-edge approach for completing three tasks on a SemEval data set using supervised machine learning [20]. Sentiment mining, categorization, and analysis face several challenges, including vast lexicons, natural language processing overheads, and bogus reviews. Numerous scholars do sentiment analysis in the languages of English, Urdu, and Arabic. In terms of accuracy, research indicates that logistic regression and SVM outperform [21]. The study on customer review categorization using sentiment analysis on hotel categorization discovered that the Lexicon model for review classification generated a 90% accuracy rate model. [22].

A paper explored Naïve Bayes to classify user sentiment using deep learning. The investigation includes many works on Hierarchical sentiment analysis using a Naive Bayes classifier. The naive Bayes classification evaluates the hierarchical sentence sentiment in hotel reviews from Traveloka [23]. Because there are so many hotel reviews on the internet that it's tough to keep track of them all, whether they are positive or negative, this study offers a solution by distinguishing positive and negative comments using the multinomial naive Bayes classifier approach [6]. The hotel review sentiment analysis, which combines the Naive Bayes algorithm and particle swarm optimization, achieves the highest accuracy of 85.00 percent [24].

Another research introduced CNN to extract local properties from a sentence's consecutive words [25]. For text sentiment analysis, CNN is also a helpful composition architecture to deal with various issues[34][35][36]. The embedding layer is utilized as a feature vectorization strategy in CNN models [26]. The Dynamic CNN method is used to handle data noise, alignment, and other data changes in sentiment analysis research. Improved results and performance of CNN graphs utilizing a dynamic k - max-pooling model and a sentiment analysis benchmark data set and test results using 5000 datasets in 500 epochs with different numbers of hidden layers and loss gradient calculations. Adam's algorithm optimization results provide the most efficient learning speed algorithms compared to others [27].

Another article describes a GRU-based RNN architecture dubbed DSWE-GRNN that is used to solve multi-class review classification problems. RNN architecture captures confidential information from particular word embeddings to train models effectively and efficiently review assessments [26]. We classify the phrases in a review as recommendations or not ideas so that reviewers don't have to comb through hundreds of reviews and can instead concentrate on actionable items and helpful suggestions [28]. The study began with creating a sentiment analysis classification model, which was then put to the test through experiments. Two categories were employed in this investigation (positive and negative). The model had a 91.9% accuracy rate, according to the results [29].

The practice of detecting and extracting information from text using natural language processing and text analysis methods is known as sentiment analysis. [30]. According to the paper's review, current learning techniques such as CNN and RNN are more efficient than classical machine learning algorithms. An article in text analysis demonstrated that CNN could identify scores and polarity reviews with 98.22% accuracy. Positive thoughts have a 95.34 percent accuracy rate, while negative ones have a 96.14 percent accuracy rate [31].

Therefore, we are proposing a model in this article for the classification of RNN user reviews. To make it simpler for hotels to identify the sentiment of hotel reviews since RNNs are very accurate at classifying sentiment.

3. Proposed Method

This paper focuses on the sentiment analysis of hotels by classifying reviews from their users. In this experiment, the sentiment classification model will organize positive and negative reviews. We define X as sentiment features and Y as a label in the training and testing process. We establish our sentiment classification model using RNN to tra in parts on datasets. The notation for RNN formulations can be seen in Table 1.

Table 1 Notation for RNN formulations

Variable	Information
x_t	Input vector (subscript t means a time sequence)
h_t	Hidden state
y	Output vector
y_t	Predicted output vector by RNN
b	Bias
f	Activation function

In this paper, we propose an RNN algorithm for sentiment analysis of hotel user review data. The RNN algorithm is a potent type of neural network and is the most promising algorithm because only those with internal memory are perfect for text data [32]. RNN uses loops inside its design to store information, which allows it to keep track of previous events. RNNs have inputs, hidden layers, and outputs.

Calculation of hidden state at a time to -t depending on the input at the time of to -t (x_t) and hidden state in the previous time (h_{t-1}). RNN equation as follows

$$h_t = f(x_t, h_{t-1}, b) \tag{1}$$

Basically, the hidden state $h_{(t)}$ is a function f from a previous hidden state $h_{(t-1)}$ and current input $x_{(t)}$. The theta is a parameter of the function f . Conceptually, equation 1 has an analogy with the full Markov chain. That is, the hidden state at the time of t depends on all hidden states and previous input.

$$\begin{aligned} h_t &= f(h_{t-1}, x_t) \tag{2} \\ &= f(x_t, f(x_t, f(x_{t-1}, h_{t-2}))) \\ &= f(x_t, f(x_{t-1}, f(\{x_t, \dots, x_{t-2}\}, \{h_1, \dots, h_{t-3}\}))) \end{aligned}$$

Softmax function to perform final classification (final output) in the form of probability in equations 3 and 4.

$$y = \text{RNN}(x_1, \dots, x_n) \tag{3}$$

$$\text{final output} = \text{softmax}(\text{MPL}(y)) \tag{4}$$

4. Experimental Setup

1. Main Idea

The basic idea of our research is to create a sentiment classification model for hotel reviews using the RNN algorithm to classify. The RNN algorithm can process sequential data and store data from the previous cell RNN recovers buried contextual information from domain-specific word insertion to effectively train the

network for review assessment classification [26]. Therefore, we explore the RNN algorithm to prepare the sentiment classification model by adopting the concept of text processing, particularly sentiment analysis.

2. Dataset

In this study, we gather hotel review datasets from Kaggle with a total of 10,000 hotel review datasets, 80% for training and 20% for testing. On our datasets, samples have positive and negative labels. Table 2 describes the dataset to be used.

Table 2 Sentiment dataset on this experiment

Dataset Label	Sentiment features	
	Training (80%)	Testing (20%)
Negative (0)	4000	1000
Positive (1)	4000	1000
Total	8000	2000

3. Data Pre-Processing

In this stage, we need to convert the feature into a vector to make it easier for the RNN algorithm to calculate the input. At the first stage of the pre-processing level, we need to undergo the case folding process. The case folding stage will convert all the letters in the document to lowercase then tokenize the text by solving a set of sentences into pieces of words. The result of tokenization will be converted to numbers or called vectors using one-hot encoding [33].

4. Classification Method

In the first phase, we get a dataset from a Kaggle site that has labels of two types: positive reviews and negative reviews. To build the model, we use two processes in running RNN algorithms, namely the training process and the testing process. To conduct our study, we gather large datasets to conduct the training process. The corpus consists of positive and negative sentiment labels with several different features. In the preprocessing stage, we undergo some procedures, including case-folding, tokenization, and vectorization (one-hot encoding). After passing through the preprocessing phase, the dataset will be converted from a sentence to a vector to enable the model to calculate the matrices in the training and testing process of the neural network.

At the final stage of the testing process, the unseen data will be fetched into our trained model to test the quality of the model. The feature will be extracted into two vector forms for the training and testing process. The standard learning model, requires a training process to build the model and the testing stage to measure the model performance in the sentiment classification of user reviews.

5. Result & Analysis

In this experiment, our proposed model can gain a trade-off by setting various hyperparameters to achieve the highest network performance. We change epoch = 30, and batch size = 64 throughout the training and testing procedure, and get an accuracy rate of 0.91 percent. Table 3 displays the results of RNN testing and training algorithms in the analysis of sentiment.

Table 3. Training, Testing on Accuracy and Loss

Hyperparameter	Optimizer	Training Accuracy	Testing Accuracy
Epoch = 30	Adam	0.9928	0.9187
Batch size = 64	RMSProp	0.9864	0.9225
Validation split = 0.2	SGD	0.6025	0.6219
LR			

Hyperparameter	Optimizer	Training Loss	Testing Loss
Epoch =30	Adam	0.0214	0.4386
Batch size = 64	RMSProp	0.0381	0.2922
Validation split = 0.2	SGD	0.6772	0.6722
LR			

To strengthen the results of training and testing on the modeling we built we also display graphs of accuracy and loss, as in Figure 1.

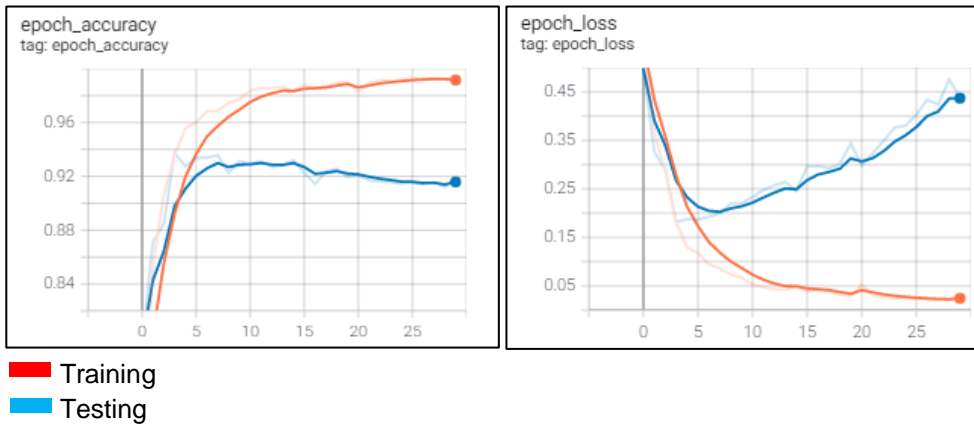


Fig 1. Accuracy chart and loss chart

This study evaluated performance models by calculating precision, recall, f1-score, and confusion matrix. The study computed accuracy as a measure of percentage certainty as positive or negative sentiment and calculated recalls determining the percentages included in a positive or negative view. We need an F1-score to compare the accuracy and recall on an average basis. Table 4 summarizes the dataset's precision, recall, and f1-score.

Table 4 Precision, recall, and f1-score

	Precision	Recall	F1-score	support
0	0.92	0.92	0.92	1014
1	0.91	0.92	0.91	986
Accuracy			0.91	2000
Macro avg	0.91	0.91	0.91	2000
Weighted avg	0.91	0.91	0.91	2000

We apply a confusion matrix to improve efficiency through RNN classification [20]. This paper also presents a tabular Confusion Matrix that depicts the model's performance on the known test data. The confusion matrix, in particular, contains information on True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) (FN). This is very helpful since the categorization results are often insufficiently represented in a single number. In Figure 2, the Confusion Matrix is shown to have the greatest TP and TN scores when the RNN method is used, with TP = 924, FP = 90, TN = 905, and FN = 81.

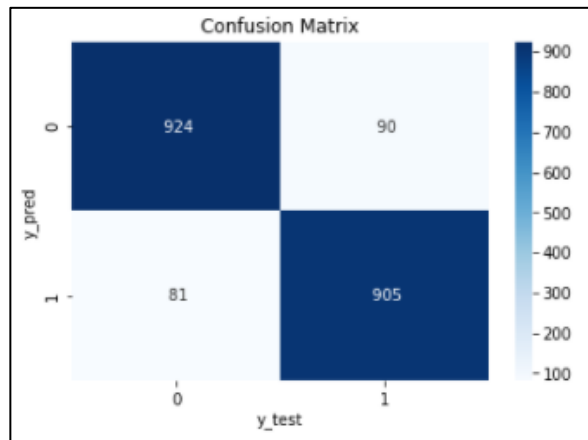


Fig. 2. Confusion Matrix

The result of this proposed confusion matrix displays data that the correct prediction by the model is 1.829 and the incorrect predicted data is 171.

6. Conclusion

Sentiment analysis in hotel reviews is a growing area of research, and websites have become a data source for knowing the quality of hotels. Hoteliers may use customer feedback to enhance their services and get better ratings. However, conventional methods to analyze unstructured textual data take a long time to determine positive sentiment or negative sentiment. Thus, we present the RNN to build the sentiment classification model to establish an effective classification model. To develop our model, we gather a huge dataset and calculate several features to obtain a promising result.

Based on the classification result, our proposed model can obtain a higher accuracy of up to 91.0%, with a precision of negative reviews reaching 92% and positive reviews worth 91%, positive and negative recall worth 92%. Positive F1 scores reached 91 %, and negative reached 92%. In addition, it produces not only excellent accuracy but also obtains a better score in evaluation metrics. Our model can get TP = 924 and TN = 905 which reflects the model to identify the sentiment effectively. Based on the results of the experiments, the model seems to be a viable option for harvesting the best accuracy outcomes. Thus, the proposed RNN can be a potential approach to deal with sentiment analysis in hotel reviews.

As future work, the classification model can develop a dynamic graph model, which could be integrated with new approaches such as the future CNN algorithm that could be combined with new techniques such as creating new architecture, including implementing GAN, GCN, or another dynamic network to improve our model.

Acknowledgment

This study is conducted under the Department of Informatics. Universitas Respati Yogyakarta, Indonesia.

References

- [1] Reza A., Riyanarto S., "Sentiment Analysis of Hotel Reviews Using Latent Dirichlet Allocation, Semantic Similarity, and LSTM," international journal of Intelligent Engineering & Systems, Vol.12, No.4, 2019.
- [2] Geetha M, Pratap S and Sumedha S., "Relationship Between Customer Sentiment and Online Customer Ratings for Hotels-An Empirical Analysis" Tourism Management, 43-54, 2017.
- [3] Dyah A, Taufiq A, et al., "Sentiment Analysis for Assessment of Hotel Services Review Using Feature Selection Approach Based-on Decision Tree." international journal of Advanced Computer Science and Application, Vol.11, No.4, 2020.
- [4] Anis, Sarah, et al., "Sentiment Analysis of Hotel Review Using Machine Learning Techniques", International Conference on Advance Intelligent Systems and Informatics Springer, Cham, 2020.
- [5] Bhargay, P. Sanjay, et al., "Sentiment Analysis for Hotel Rating using Machine Learning Algorithms." International Journal of Innovative Technology and Exploring Engineering (IJITEE) 8.6, 2019.
- [6] Arif A F, Yuliant S and Said A., "Sentiment analysis on hotel reviews using Multinomial Naïve Bayes classifier" Journal of Physics: Conference Series, 1192, 012024, 2019.
- [7] Chang, Victor, et al., "An improved model for sentiment analysis on luxury hotel review," Expert Systems: e12580. 2020.
- [8] Demir, Yunus Emre, et al., "Sentiment Analysis for Hotel Attributes from Online Reviews," Innovations in Intelligent Systems and Applications Conference (ASYU). IEEE, 2020.
- [9] Rybakov, Valery, and Alexey Malafeev., "Aspect-based sentiment analysis of Russian hotel reviews," CEUR Workshop Proceedings, 2018.
- [10] Zvarevashe, Kudakwashe, and Oludayo O. Olugbara., "A framework for sentiment analysis with opinion mining of hotel reviews," Conference on information communications technology and society (ICTAS), IEEE, 2018.
- [11] Nie, Ru-xin, et al., "Hotel Selection Driven by Online Textual Reviews: Applying a Semantic Partitioned Sentiment Dictionary and Evidence Theory," International Journal of Hospitality Management 88, 102495, 2020.
- [12] Park, Eunil, et al., "Understanding customers' hotel revisiting behaviour: a sentiment analysis of online feedback reviews," Current Issues in Tourism, Vol. 23, No.5, pp. 605-611, 2020.
- [13] Miharja, Jaja, Jordy L P, and Hardianto., "Comparison of Machine Learning Classification Algorithm on Hotel Review Sentiment Analysis (case study: Luminor Hotel Pecenongan), " Pilar Nusa Mandiri, Vol.6, No.1, pp. 59-64, 2020.
- [14] Zvarevashe K, Olugbara O.O., "A framework for sentiment analysis with opinion mining of hotel reviews," Conference on Information Communications Technology and Society, ICTAS – Proceedings, 2018.
- [15] Nasichuddin, Moch A, Teguh B and Widyawan., "Performance Improvement Using CNN for Sentiment analysis," IJITEE (International Journal of Information Technology and Electrical Engineering), Vol. 2, No.1, pp 9-14, 2018.
- [16] Basiri, M Ehsan et al., "ÁBCDM: an Attention-based Bidirectional CNN-RNN Deep Model for Sentiment Analysis," Future Generation Computer System, Vol.115, pp. 279-294, 2021.
- [17] Ay Karakus, Betul, et al., "Evaluating Deep Learning Models for Sentiment Classification," Concurrency and Computation: Praticce and Experience, Vol. 30, No. 21, e4783, 2018.
- [18] Seo, Seungwan, et al., "Coparative Study of Deep Learning-Based Sentiment Classification," IEEE Access, Vol. 8, pp. 6861-6875, 2020.
- [19] Aurangzeb, khursheed, Nasir A, and Musaed A., " Aspect Based Multi-Labeling Using SVM Based Ensembler," IEEE Access, Vol. 9, pp. 26026-26040, 2021

- [20] Abro, Sindhu, et al., "Aspect Based Sentimental Analysis of Hotel Reviews: a Comparative Study," *Sukkur IBA Journal of Computing and Mathematical Sciences*, Vol. 4, No. 1, pp. 11-20, 2020.
- [21] Nazir, Kashif M, et al., "Sentiment Analysis of User Reviews about Hotel in Roman Urdu," 14th International Conference on Open Source Systems and Technologies (ICOSST), IEEE, 2020.
- [22] Abbas, Muhammad, et al., "Multinomial Naive Bayes Classification Model for Sentiment Analysis," *IJCSNS*, 2019.
- [23] Kurniawan, Sandy, Retno K, and Melny E T., "Hierarchical Sentence Sentiment Analysis Of Hotel Reviews Using The Naïve Bayes Classifier," 2nd International Conference on Informatics and Computational Sciences (ICICoS) IEEE, 2018
- [24] Aji, Sopian, et al., "Review Sentiment Analysis of World Class Hotel Using Naive Bayes Classifier And Particle Swarm Optimization Method," pp. 554-560, 2018.
- [25] Yan, Cheng, et al., "Text Sentiment Orientation Analysis of Multi-Channels CNN and BiGRU Based on Attention Mechanism," *Journal of Computer Research and Development*, Vol. 57, No.12, pp. 2583, 2020.
- [26] Hassan J, Shoaib U., "Multi-class Review Rating Classification Using Deep Recurrent Neural Network," *Neural Processing Letters*, pp. 1031-1048, 2020.
- [27] Wanda, P and Huang Jinjie., "Model of Sentiment Analysis with Deep Learning in Social Network Environment," *IEEE 2nd International Conference on Electronic Information and Communication Technology*, 2019.
- [28] Ezen-can, Aysu, and Ethem F C., "Hybrid RNN at SemEval-2019 task 9: blending information sources for domain-independent suggestion mining," *Proceedings of the 13th International Workshop on Semantic Evaluation*, 2019.
- [29] Kurniasari, Lilis, and Arif Setyanto., "Sentiment Analysis using Recurrent Neural Network," *Journal of Physics: Conference Series*, Vol. 1471, No. 1, 2020.
- [30] Arora H and Mamta B., "Feature Extraction Through Sentiment Analysis of Tourist Sentiment Using Deep Learning Techniques Like CNN, RNN and LSTM," *International Journal of Recent Technology and Engineering*, 2020.
- [31] Lal k and Mishra N., "Feature Based Opinion Mining on Hotel Reviews Using Deep Learning," *Lecture Notes on Data Engineering and Communications Technologies*, 2020.
- [32] Kanagachidambaresan G.R, et al., "Recurrent Neural Network," *Springer Innovations in Communication and Computing*, 2021.
- [33] Hongyu Liu, et al., "CNN and RNN based payload classification methods for attack detection," *Accepted Manuscript*, 2018.
- [34] Wanda, P. and H. Jie. "DeepSentiment: Finding Malicious Sentiment in Online Social Network based on Dynamic Deep Learning.", 2019.
- [35] Wijaya, Nurhadi., "Evaluation of Naïve Bayes and Chi-Square performance for Classification of Occupancy House," *International Journal of Informatics and Computation*, [SI], v. 1, n. 2, p. 46-54, Feb. 2020.
- [36] Diqi, Mohammad, Muhdalifah, Mujastia., "Design and Building Javanese Script Classification in The State Museum of Sonobudoyo Yogyakarta," *International Journal of Informatics and Computation*, [SI], v. 1, n. 2, p. 35-45, Feb. 2020.