




**LONG SHORT-TERM MEMORY APPROACH  
IN SENTIMENT ANALYSIS USING YELP DATASET**

*FINAL PROJECT*



Abi Fauzan  
41516020002

**DEPARTMENT OF INFORMATICS  
FACULTY OF COMPUTER SCIENCE  
UNIVERSITAS MERCU BUANA**

**JAKARTA  
2020**

MERCU BUANA



**LONG SHORT-TERM MEMORY APPROACH  
IN SENTIMENT ANALYSIS USING YELP DATASET**

*Final Project*

Submitted to Complete Terms  
Acquired a Computer Bachelor Degree

Created By:  
Abi Fauzan  
41516020002

U DEPARTMENT OF INFORMATICS  
FACULTY OF COMPUTER SCIENCE  
UNIVERSITAS MERCU BUANA  
JAKARTA

2020

## ORIGINALITY STATEMENT SHEET

The undersigned below:

Student Number : 41516020002

Student Name : Abi Fauzan

Title : Long Short-Term Memory Approach in Sentiment Analysis  
using Yelp Dataset

Stating that my Final Project Report is the work of my own and not plagiarism. If it is found in my Final Project Report that there is an element of plagiarism, then I am ready to get academic sanctions related to it.

Jakarta, June 17<sup>th</sup> 2021



METERAI  
TEMPEL  
C29E1AJX273678034

Abi Fauzan



UNIVERSITAS  
MERCU BUANA

## FINAL PROJECT PUBLICATION STATEMENT

As a Universitas Mercu Buana student, I undersigned below:

Student Number : 41516020002  
Student Name : Abi Fauzan  
Title : Long Short-Term Memory Approach in  
Sentiment Analysis using Yelp Dataset

By giving permission and approval of Non-exclusive Royalty Free Right to Universitas Mercu Buana for my scientific work entitled above along with available devices (if necessary).

With this **Non-exclusive Royalty right**, Universitas Mercu Buana has the right to store, transfer / format, manage in the form of a database, administer and publish my final project.

Furthermore, for the sake of scientific development in Universitas Mercu Buana environment, I give permission to Researchers in the Research Lab of Computer Science Faculty, Universitas Mercu Buana to use and develop the existing result of the research of my final project for the research and publication purposes as long as my name is stated as author/creator and Copyright owner.

Hereby I made this statement in truth.

UNIVERSITAS  
MERCU BUANA

Jakarta, 17<sup>th</sup> June 2021



Abi Fauzan

## STATEMENT LETTER OF FINAL PROJECT OUTPUT

As a Universitas Mercu Buana student, I undersigned:

Student Number : 41516020002

Student Name : Abi Fauzan

Title : Long Short-Term Memory Approach in Sentiment Analysis using Yelp Dataset

Declare that my Final Project Output is as follows:

No	Luaran	Jenis	Status
1	Scientific Publications	National Journal of Non-Accreditation	Submitted
		Accredited National Journal	
		International Journal with No Reputation	Accepted
		Reputable International Journal	
Submitted/Accepted:		Journal Name :	
		ISSN :	
2	Working Sheet, is the result of research material as a complete Journal Article. Consist of (minimum 4)	Literature Review	[ ]
		Analyst Result & Develop Application	[ ]
		Source code	[ ]
		Data set	[ ]
		Experiment Process	[ ]
		All Experiment Result	[ ]
		....	
3	HAKI Submitted / Registered	HKI	Submitted
		Patent	Recorded
		No & Date Application	
		No & Date Recorded	

This statement I made in truth.

Jakarta, 17<sup>th</sup> June 2021




Abi Fauzan

## VALIDITY SHEET

NIM : 41516020002  
Name : Abi Fauzan  
Final Project Title : Long Short-Term Memory Approach in Sentiment  
Analysis using Yelp Dataset

This Final Project has been examined and defended as one of the requirements for obtaining a Bachelor's degree in the Informatics Engineering Study Program, Faculty of Computer Science, Universitas Mercu Buana.

Jakarta, 15 September 2020



(Dr. Leonard Goeimanto)

## VALIDITY SHEET

NIM : 41516020002  
Name : Abi Fauzan  
Final Project Title : Long Short-Term Memory Approach in Sentiment Analysis using Yelp Dataset

This Final Project has been examined and defended as one of the requirements for obtaining a Bachelor's degree in the Informatics Engineering Study Program, Faculty of Computer Science, Universitas Mercu Buana.

Jakarta, 31 Agustus 2020



UNIVERSITAS  
( Drs. Achmad Kodar, MT )  
MERCU BUANA



## VALIDITY SHEET

NIM : 41516020002  
Name : Abi Fauzan  
Final Project Title : Long Short-Term Memory Approach in Sentiment Analysis  
using Yelp Dataset

This Final Project has been examined and defenced as one of the requirements for obtaining a Bachelor's degree in the Informatics Engineering Study Program, Faculty of Computer Science, Universitas Mercu Buana.

Jakarta, 31 Agustus 2020



(Dr. Ida Nurhaida, MT)



## VALIDITY SHEET

Student Number : 41516020002

Student Name : Abi Fauzan

Title : Long Short-Term Memory Approach in Sentiment  
Analysis using Yelp Dataset

This Final Project has been examined and defended as one of the requirements for obtaining a Bachelor's degree in the Informatics Engineering Study Program, Faculty of Computer Science, Universitas Mercu Buana.

Jakarta, 27<sup>th</sup> August 2020


Approved,




(Dr. Mujiono Sadikin, MT)  
Supervisor

Acknowledged,

UNIVERSITAS  
MERCU BUANA



(Diky Firdaus, S.Kom, MM)  
Informatics Thesis Coordinator



(Desi Ramavanti, S.Kom, MT)  
Head of Informatics Department

## ABSTRAK

Nama : Abi Fauzan  
NIM : 41516020002  
Pembimbing TA : Dr. Mujiono Sadikin, MT.  
Judul : Long Short-Term Memory Approach in Sentiment Analysis using Yelp Dataset

Algoritma *deep learning* telah digunakan untuk mencapai hasil yang luar biasa dalam aplikasi *Natural Language Processing (NLP)*. Analisa Sentimen adalah salah satu bagian dari aplikasi NLP yang mengekstraksi informasi emosional dari sebuah teks. Penelitian ini bertujuan untuk menghasilkan model *deep learning* yang dapat memprediksi data review restaurant menggunakan *Multi-Layer Perceptron (MLP)* dan *Long Short-Term Memory (LSTM)*. Dataset dilatih menggunakan *Word2vec* untuk menghasilkan representasi kata berupa vektor yang akan digunakan sebagai input pada model *deep learning*. Dalam proses eksperimen, model MLP akan dibandingkan dengan model LSTM untuk mendapatkan akurasi terbaik. Berdasarkan hasil percobaan menunjukkan bahwa model LSTM mengungguli model MLP dengan nilai akurasi 91% dibandingkan dengan hasil model MLP dengan nilai akurasi 76%.

**Kata kunci:** *LSTM, MLP, NLP, sentiment analysis, word embeddings*

## ABSTRACT

Name : Abi Fauzan  
Student Number : 41516020002  
Counsellor : Dr. Mujiono Sadikin, M.T.  
Title : Long Short-Term Memory Approach in  
Sentiment Analysis using Yelp Dataset

Deep learning algorithms have been used to achieve great results in Natural Language Processing (NLP) applications. Sentiment Analysis is a part of NLP application that extracts emotional information from the texts. This research aims to generate deep learning models which can predict a particular restaurant review using a Multi-Layer Perceptron (MLP) and Long Short-Term Memory (LSTM) Neural Network. The dataset used in this research is a restaurant review taken from Yelp website. The dataset is trained using Word2vec word embedding to generate word vector representation which will be used as deep learning input. In the experiment process, the MLP model will be compared with the LSTM model to get the best accuracy. Based on the experiment results shows that the LSTM model is outperformed the MLP model with 91% accuracy compared to the result of MLP model with 76% accuracy.

**Keywords:** *LSTM, MLP, NLP, sentiment analysis, word embeddings*

## ACKNOWLEDGEMENT

Alhamdulillah, praise be to Allah the Almighty and the most Merciful that because of His blessings I could finally finish my thesis to attain a Bachelor's degree (S1) on Informatics Department of Universitas Mercu Buana. In this occasion, the author would like to express deepest gratitude to all who helped me in completing the writing of this thesis to:

1. Dr. Mujiono Sadikin, MT. as a thesis supervisor who guided the author's thesis and gives suggestion for author so it can finish on time
2. Mrs. Desi Ramayanti, S.Kom, MT. as the head of Informatics Department of Universitas Mercu Buana.
3. Dr. Mujiono Sadikin, MT. as Dean Faculty of Computer Science of Universitas Mercu Buana
4. Mrs. Prastika Indriyanti, S.Kom, MCS. as the secretary of Informatics International Class
5. Mr. Diky Firdaus, S.Kom, MM. as the Thesis Coordinator of Informatics Department
6. My beloved Parents and my whole family who have prayed, motivated, encourage, reminded the author to finish the college as soon as possible
7. Colleagues from International Class of Universitas Mercu Buana, especially classmates from Informatics International Class 2016
8. Furthermore, the author would like to say thanks to many people who helped in the process of writing this thesis report that couldn't be mention one by one.

The final word the author apologizes for any mistakes that have been made. The author realizes that this thesis is not perfect yet, therefore the author will be glad to receive any suggestion and criticisms which is built for the sake of perfection. Hopefully this thesis can provide benefits to encourage further research. Ameen.

Jakarta, June 3<sup>rd</sup>, 2020

Abi Fauzan

## TABLE OF CONTENT

COVER PAGE.....	i
TITLE PAGE .....	i
ORIGINALITY STATEMENT SHEET .....	ii
FINAL PROJECT PUBLICATION STATEMENT .....	iii
STATEMENT LETTER OF FINAL PROJECT OUTPUT .....	iv
EXAMINER APPROVAL SHEET.....	v
VALIDITY SHEET .....	viii
ABSTRAK.....	ix
ABSTRACT.....	x
ACKNOWLEDGEMENT .....	xi
TABLE OF CONTENT .....	xii
JOURNAL .....	1
WORKING SHEET .....	11
PART 1. LITERATURE REVIEW .....	12
PART 2. ANALYSIS AND DESIGN .....	14
PART 3. SOURCE CODE .....	15
PART 4. DATASET.....	22
PART 5. EXPERIMENT PROCESS.....	23
PART 6. EXPERIMENT RESULT .....	28
REFERENCES .....	30
LEGAL DOCUMENT ATTACHMENT .....	32
CORRESPONDENCE ATTACHMENT .....	34

## JOURNAL

# LONG SHORT-TERM MEMORY APPROACH IN SENTIMENT ANALYSIS USING YELP DATASET

Abi Fauzan<sup>1</sup>, Mujiono Sadikin<sup>1</sup>

<sup>1</sup>Faculty of Computer Science, Universitas Mercu Buana, Jakarta 11650, Indonesia  
Email: 41516020002@student.mercubuana.ac.id

### **ABSTRACT**

Deep learning algorithms have been used to achieve great results in Natural Language Processing (NLP) applications. Sentiment Analysis is a part of NLP application that extracts emotional information from the texts. This research aims to generate deep learning models which can predict a particular restaurant review using a Multi-Layer Perceptron (MLP) and Long Short-Term Memory (LSTM) neural network. The dataset used in this research is a restaurant review taken from Yelp website. The dataset is trained using Word2vec word embedding to generate word vector representation which will be used as deep learning input. In the experiment process, the MLP model will be compared with the LSTM model to get the best accuracy. Based on the experiment results shows that the LSTM model is outperformed the MLP model with 91% accuracy compared to the result of MLP model with 76% accuracy.

**Keywords:** *deep learning, LSTM, MLP, NLP, sentiment analysis, word embeddings.*

### **1.0 Introduction**

Billions of data generated in various sources from the internet every day. Text-based data can be found on social media services, forums, articles, and sites containing public opinion on products, brands, services, politics, or any other topic. The data is then used by companies, governments, and other fields to create marketing analyzes, product reviews, and community services. Thus, the input from users or consumers will be obtained efficiently. Sentiment Analysis (SA) is one of the opinion mining study, that is process understanding and transforming unstructured data into structured data automatically to get an understanding of the information contained in an opinion sentence [1]. The essence of SA system is finding the meaning in a text or document contained opinion or sentiment. In a single sentence, it is occasionally containing some opinions or facts in unstructured conditions [2]. Therefore, it is required to transform this kind of sentence to a structured condition. SA helps in determining opinion as positive or negative from the text towards which sentiment is directed [3]. One of the necessary processes in the SA system is to extract word into a numerical representation namely feature extraction, in which computers can process it. In general, feature extraction techniques are Bag of Word, TF-IDF, and Word Embedding (Word2vec, GloVe) [4], [5]. These techniques are useful to explore the

potential information on the datasets or dictionary and represent original samples as feature vectors which will be used as input to the machine learning methods later. Word embedding with Word2vec technique is used in this research.

Many researchers have shown the effectiveness of deep learning techniques in sentiment analysis topic in order to extract meaningful information expressed by the people. MLP is a type of feedforward deep learning network. It has achieved great result in sentiment analysis task. Another type of deep learning network that can solve sentiment analysis task is LSTM. Text review is a sequence data, LSTM network take advantages of this particular problem since it learns context in chronological order from each text in dataset. In addition, the case of a long review text, thanks to additional layer in LSTM cell namely memory cell which will decide what information will be passed or removed.

Thus, the research question that appear is how to propose an effective deep learning models which can make a prediction on a given new particular restaurant review data. To achieve this question, a comparison of MLP model and LSTM model is proposed, which is associated with Word2vec word embedding. It has shown that LSTM model outperformed the MLP model. The strong idea of LSTM model that it allows to control the text flow over the long text. For validation process, both models are compared with different hyperparameter configuration to conclude the research study. Splitting the dataset into train set and test set is the major idea to validate the result. Thus, the LSTM model is the one that achieved the highest accuracy of 91 %.

In this proposed research, SA is performed to classify the polarity of given restaurant reviews whether it is a positive review or negative review which is known as a binary text classification task. The dataset used in this research is Yelp dataset which contains more than 6 million of restaurant reviews in the USA. Each instant of dataset has 9 attributes i.e. review id, user id, business id, stars, date, text, useful, funny, and cool [6] The attributes used for this research are only text (contain a review from a user) and stars (contain rating of the restaurant). Due to high computational process and education purposes, the dataset is truncated to only 100.000 reviews which contain 50.000 score rating above 3 and 50.000 score rating below 3. Feature vectorization is done by using Word2vec word embedding technique [7]. In this experiment, Recurrent Neural Network (RNN) particularly LSTM network is implemented as a deep learning with hyperparameters optimization such as differentiate of batch number and the implementation of dropout regularization for validating the generated models.



## 2.0 Related Study

Research by using Yelp dataset is conducted by Fan and Khademi in [8] on the restaurant's average star rating prediction. They applied the unigram model combined with Part-of-Speech (POS) tagging as a feature engineering, Support Vector Machine (SVM), and decision tree were applied for prediction. The dataset used in their experiment consists of 4234 list of restaurants and 35465 reviews. Their study relates to this proposed research in the sense of NLP task that is opinion mining. In term of feature extraction process, the implementation of POS tagging is indeed produced good feature vector, but this does not include the semantic meaning of the data. Thus, compared to the proposed research, the deep learning models are associated with Word2vec word embedding which generated feature vector with meaningful representation.

Miedema in [9] proposed an advanced research that is used to classify tweets data. They used LSTM model to do classification tasks which often use in long-term sequence data. The result of their research achieved the accuracy of 86% correctly classified. There was no pre-processing phase as they used a pre-build IMBD dataset that already vectorized in number format. Furthermore, performing hyperparameter optimization on their model parameters could improve the result of the accuracy.

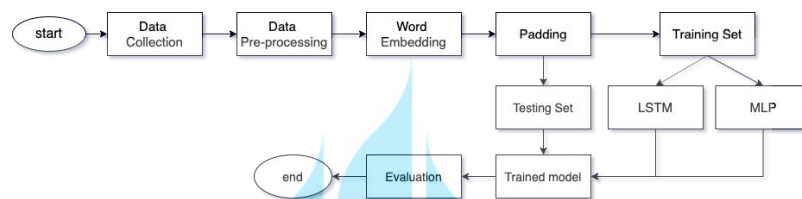
A comparison of classification methods in extracting drug name entity in the medical text is proposed by Sadikin, *et.al* in [10]. They proposed a new representation technique based on the characteristics of word distribution and word similarities to overcome some issues in the medical text mining, such as multiple names for the same drug, the lack of labeled dataset sources in medical domain and external knowledge. In their research, three data representation techniques were conducted, the first is with standard neural network model, that is MLP, the second techniques were Deep Believe Network (DBN) and Sparse Autoencoder (SAE). The last technique was using LSTM model. The results of all models have given the best result to LSTM model compared to the state of the art, with its average f-score was 86%. Thus, LSTM model gives the best result of the sequence data representation for the sentences of medical text.

Research conducted by Wiranda and Sadikin in [11] was predicting product sales in the pharmacy industry. The main topic of their research is the ability to predict the direction of prices of the company's product sales. In their research, time-series prediction is done using LSTM model. They applied performance parameters of Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE) thought the dataset. Thus, the results of

the model evaluation have shown that the use of LSTM network in predicting sales was 13,762,154.00 for RMSE in rupiah values, and 12% for MAPE. In addition, their research has proved that LSTM model achieve great result on both time-series data and text data since they contain sequence information in the dataset.

### 3.0 Methodology

This study aims to analyze the reviews and classify the sentiment based on the reviews available. The methodology used in this study consists of five main steps, they are data collection, data pre-processing, sentence conversion, classification model, and evaluation. The research stage is described in figure 1.



**Figure 1** Research stage

The first process of the research stage is data collection and data exploration. The restaurant review data was taken from the Yelp website, [www.yelp.com/dataset](http://www.yelp.com/dataset). A binary classification is performed where each review contains positive or negative sentiment [3]. Data pre-processing includes stop word and punctuation removal, case folding, and stemming. The next process is the conversion of the cleaned dataset into data that were ready to be used in the classification model as the input [12]. This process includes making a dictionary words from the dataset and converting it into numbers that already padded. Vector conversion is conducted in which the values are taken from the training dataset of the restaurant review by using Word2vec word embedding. A sequential model LSTM is proposed to recognize the sentiment analysis of the restaurant reviews. LSTM network is an advanced version of RNN that is mostly used to process long sequential data such as SA [13]. The final process is to evaluate the accuracy of the model for restaurant reviews. The evaluation included the calculation of the confusion matrix.

### 3.1 Data collection

In this study, a restaurant review of Yelp Dataset is used. Yelp Dataset is a collection of user data, reviews, and businesses in a large-scale dataset provided by Yelp company that is used for individual, educational and academic purposes [6]. The review datasets consist of more than six million reviews of data generated by the user. Due to high computational deep learning training process, the review data is reduced to 100.000 reviews data which

consists of attributes text (review data) and stars (rating of the restaurant). Each review is converted into binary labels 0 and 1 in which rating 1-2 grades as negative and 4-5 grades as positive. Table 1 shows a sample of data review which will be processed in a further step.

**Table 1** Sample data

Text	Stars	Sentiment
The best in Wisconsin and being the dairy state that's saying a lot The ice cream is a great consistency and there are so many creative flavors like cake batter fudge zanziberry shortcake...	5	pos
Worst experience ever!  Completely unorganized, unfriendly staff and the food was cold and bland.... extremely disappointing	1	neg

### 3.2 Data pre-processing

Pre-processing is done by using The scikit-learn's train\_test\_split function as the first process to split the review data into an 80% train set and 20% test set. To embed the dataset into the Word2vec model, the data is processed with a function that returns a list of parsed sentences with integer values. For each sentence, noise removal is done by removing punctuation, numbers, and HTML tag, stop word removal, and case folding. Table 2 shows the result of the pre-processing of the review data.

**Table 2** Pre-processing phase

Pre-processing	Sentence
	It's an adorable old home converted into a cafe/bar/restaurant.
Case folding	it's an adorable old home converted into a cafe/bar/restaurant.
Filtering	its an adorable old home converted into a café bar restaurant.
Tokenization	[its] [an] [adorable] [old] [home] [converted] [into] [a] [café] [bar] [restaurant]

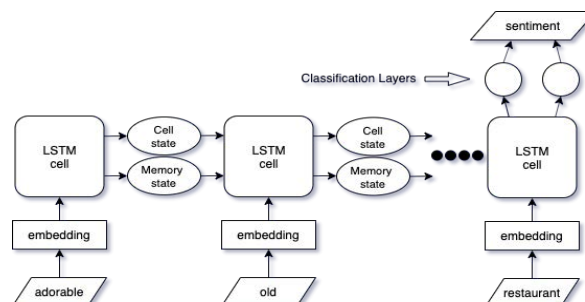
Stop word removal	[adorable] [old] [home] [converted] [café] [bar] [restaurant]
-------------------	--

### 3.3 Sentence conversion

Sentence conversion is processed using Word2vec model after the pre-processing step is done. A vocabulary is built by embedding training data into Word2vec model. This vector conversion learns and sets the vector representation of each word [14]. By default, the training algorithm in Word2vec is CBOW. There are two algorithms provided by Word2vec model, they are CBOW and Skip-gram [15]. CBOW algorithm is a process of calculating the current word from the context word. It is similar to maximize the probability of the target word by looking at the context. Skip-gram algorithm is a process of calculating the context of word from the word at the current position in the sentence. The Embedding Word vector will produce a padded list of sentences as classification features. The implementation of using Word2vec as a word embedding technique can help to improve the classification model since the algorithm of Word2vec is to find similar words that will have similar vectors [7].

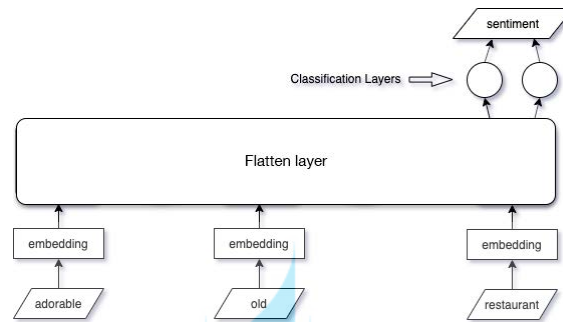
### 3.4 Classification model

The main idea of the classification model is to train the LSTM network and compared it with MLP model [16] through text review data. On the other hand, MLP can be described as feedforward NN which does not have deep layer, thus the information flow in a simple direction. Tokenization proceeds in each review that produced a list of parsed sentences and is generated a vocabulary of words. The sentiment classification is being done by feeding the word embedding obtained through a vocabulary index-based into the LSTM network. The proposed LSTM model is illustrated in figure 2.



**Figure 2** Proposed LSTM architecture

Regularization technique such as Dropout is applied in the LSTM model to prevent the model from being overfitting, it could affect the performance of LSTM model [17]. In the process of training the model, the use of Adaptive Moment Estimation (Adam) [18] optimization is implemented for updating the weights and bias values in each neuron across iterations. To predict the outcomes of sentiment, Sigmoid activation function is used to scale the output data points between 0 and 1 [19].



**Figure 3** MLP architecture

From the above architecture in figure 3, a Flatten layer is applied after the embedding process. This layer reshapes the dimension of the input vector equal to one-dimension array of elements. Thus, the flattened layer from the input is classified in a Dense layer with two nodes and activated with sigmoid function.

### 3.5 Evaluation model

In this study, the validation process includes comparison of different parameters for performing sentiment analysis. Also, two different models are compared (LSTM model and MLP model) to get a clearer result. The accuracy is calculated based on 5 iterations, between the values of batch size 32, 64, and 128 and using the ADAM optimizer. Table 3 shows the model configurations.

**Table 3** Model configuration

<b>Word2vec dimension</b>	300
<b>Epochs</b>	5
<b>Batch size</b>	[32, 64, 128]
<b>Optimizer</b>	Adam
<b>Regularization</b>	Dropout (0.2)

In this study, the performance of sentiment classification is evaluated by using given formulas calculated as the following Eq. (7) – (9).

$$\text{Accuracy} = ( TP+TN ) / ( TP+TN +FP+FN ) \quad (7)$$

$$\text{Precision} = TP / ( TP + FP ) \quad (8)$$

$$\text{Recall} = TP / ( TP+FN ) \quad (9)$$

The confusion matrix is shown in table 4 to evaluate the classification result of LSTM model and MLP model.

**Table 4** Confusion matrix formulation

	Actual Positive	Actual Negative
Predict Positive	TP	FP
Predict Negative	FN	TN

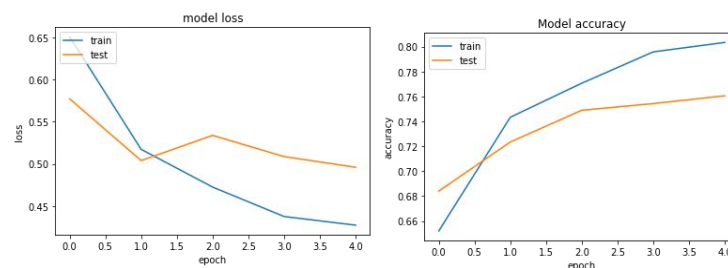
#### 4.0 Experiment result and analysis

The experiment is done by using 100.000 restaurant reviews from Yelp dataset. There are 50.000 positive reviews and 50.000 negative reviews, thus, the class distribution is balance. The training set is defined as 80% of the data and the testing set is defined as 20% of the data. And the training is split to get validation data as 20%.

**Table 5** Accuracy results of MLP model.

Model	Epochs	Batch size	Accuracy
Model 1	5	32	76%
Model 2	5	64	75%
Model 3	5	128	74%

The result of MLP model shows in table 5 that by using a batch size of 32 gives the highest result of 76% in model 1. A smaller batch size has shown performs better than a larger batch size, and this will give to a higher accuracy result in neural network. Figure 4 shows the model 1 MLP accuracy and loss result.



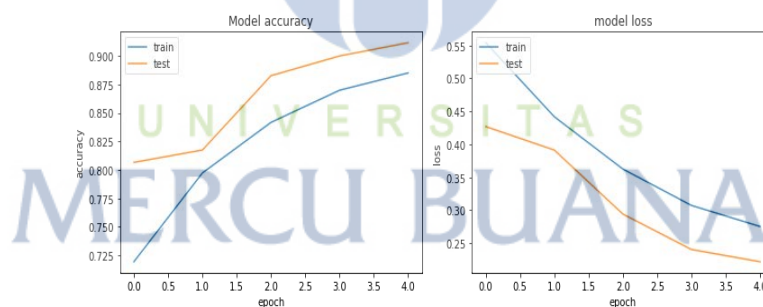
**Figure 4** MLP model 1 accuracy and loss

According to the design of a multi-layer perceptron, each input act independently, in other words, each input in MLP model has no relation from others input. Thus, this kind of model is not efficiently performed classification in sequential data. Also, based on the above loss result, the test loss is higher than training loss at second iterations which caused overfitting in neural network model. To overcome this issue, the LSTM model which learns better on sequential data is performed to compare the result of accuracy.

**Table 6** Accuracy results of LSTM model.

Model	Epochs	Batch size	Accuracy
Model 1	5	32	<b>91%</b>
Model 2	5	64	90%
Model 3	5	128	87%

It has been examined the performance model of LSTM network with Word2vec word embedding with different sets of parameters. Model 1 in table 6 achieved the best accuracy result of 91% with 32 batch size value. Based on the above table, the LSTM model is learned better than MLP, this is simply because, in LSTM cell, there is a hidden state called memory cell, that captured context from the input and determine which value to process or eliminate in the next calculation. Figure 5 shows the LSTM model accuracy and loss.



**Figure 5** LSTM model 1 accuracy and loss

After analyzing the result of MLP and LSTM model on the testing set, the value of confusion matrix, Accuracy, Precision, Recall, and F-Score are obtained based on model 1 in each model in the table 7 and table 8.

**Table 7** MLP model 1 calculated class prediction

MLP Model 1	Actual Positive	Actual Negative
Predict Positive	6969	3024



Predict Negative	1745	8262
------------------	------	------

In MLP model, there are still many incorrect class predictions compared to LSTM model. This issue occurs according to the training process of MLP which uses the basic implementation of a multi-layer perceptron where each input has no relation from previous hidden state or previous input.

**Table 8** LSTM model 1 calculated class prediction

LSTM Model 1	Actual Positive	Actual Negative
Predict Positive	8944	1049
Predict Negative	828	9179

Thus, in LSTM model, the incorrect class prediction is reduced by almost 20% based on the use of gating system in the training process of LSTM model.

## 5.0 Conclusion and future work

According to the result of the experiment, it can be concluded that the LSTM model is outperformed the MLP model, due to the control text flow in the LSTM cell so it can learn context efficiently from the input data. Thus, it is a good solution for classification task. Furthermore, the implementation of different batch size values will validate the generated model through the dataset. From this experiment, by using batch size 32 towards the LSTM model could give the highest accuracy at 91%. In this case, higher batches value not guaranteed that the accuracy result will be as good as the smaller ones. Thus, with 91% accuracy, the use of LSTM network combined with Word2vec word embedding and several hyperparameters tuning could be used in the case of sentiment classification particularly on classify the restaurant review on Yelp dataset. In future work, the proposed model can be implemented for other datasets, and determine the polarity based on the sequence data such as text review. The implementation of stacked LSTM and Bidirectional LSTM could improve the performance of the proposed study. In addition, the proposed study can be improved by exploring other word embeddings such as GloVe and Fasttext [20], [21].

## WORKING SHEET

This working sheet aims to understand how the process of Sentiment Analysis calculation by implementing the LSTM model and to test the accuracy of the model with restaurant review dataset. In this section, further analysis has been described regarding the results and materials of the final project research. There are literature review, analysis of the project, source code, description of the dataset used in this research, experiment process, and experiment results.

