Chatbot for Pharmacy

SIM JIA YAO
Faculty of Information Science and Technology
Multimedia University Melaka
Melaka , Malaysia
1171303327@student.mmu.edu.my

TAN KIAN LONG
Faculty of Information Science and Technology
Multimedia University Melaka
Melaka , Malaysia
1181300023@student.mmu.edu.my

CHEE YONG EN
Faculty of Information Science and Technology
Multimedia University Melaka
Melaka , Malaysia
1181300217@student.mmu.edu.my

NG ZI XUAN
Faculty of Information Science and Technology
Multimedia University Melaka
Melaka , Malaysia
1181300765@student.mmu.edu.my

***Abstract*—** In this COVID-19 pandemic period, it is a must to adhere to the Standard Operating Procedure (SOP) set by government in order to improve the situation. One of the most efficient ways to cut down the chain of COVID-19 infection is to stay at home and avoid being in crowded areas. Hence in this research, a chatbot for pharmacy is created to provide immediate assistance to the customer's request without having them travelling to pharmacy. In this research, Natural Language Processing technique is used to identify and recognize the input words and two models : Convolutional Network (CNN) and Multilayer Perceptron (MLP) are used for training the models. The results are then compared to determine which method is more efficient

Keywords— Chatbot, Pharmacy, NLP, CNN, MLP

I. INTRODUCTION

In the era of globalisation, many things have been automated ranging from the production in factories to the customer service. This has brought a lot of convenience to us as many of these automated processes have fewer errors and tend to be more efficient. One of the most convenient automated inventions is chatbot. A chatbot is a software application that is used to conduct a conversation via text or text-to-speech with a live human. Chatbots are often used in customer service, request routing and information gathering. In the chatbot system, natural language processing technique is often used to recognize, identify and extract the keywords and give responses accordingly. In this research, we are applying this simple chatbot system to pharmacies in order to provide some help and convenience to the patients as well as the medical workers. This is due to the recent outbreak of COVID-19 pandemic and the government has encouraged the citizens to stay at home in order to break the chain of infection. Hence with the help of the chatbot system, the patients that have mild or light discomfort are no longer needed to visit the pharmacy when their problems can be solved by the chatbot. If their problems persist or worsened, they will then be assisted by live pharmacists to help solve their problems. Therefore, this chatbot system is designed to help the citizens to abide by the rules of movement control order (MCO) set by Malaysian Government.

II. RELATED WORK

A. *Combating Depression in Students using an Intelligent ChatBot: A Cognitive Behavioral Therapy.*

Based on the paper, [1] have proposed an intelligent chatbot which able to determine user's emotion and mentality after communicate with them. The experiment consists of two important part which are training and testing and classifying. To train this model, the ISEAR dataset has been in the training and testing phase. In the training and testing phase, the dataset has been tokenized into a huge collection of dictionaries by using Punkt Sentence Tokenizer. After that, to form word vector of corpus, the authors are using GloVe with 100 dimensions as the word vector representation. Then, the word vector is will be mapped into word embedding. Finally, it will use the classifier to train it. In this research, the authors used three types of deep learning algorithms which are CNN, RNN and HAN. After training the model, they get the user's chat text as input to determine the user's emotion. It consists of five labels of emotion such as happy, joy, shame, anger and disgust. According to different labels, the system will generate different advice of user.

Advantages: It able to help the user to check their emotion so that they can discover it and treat it as soon as possible.

Disadvantages: Limitation of the dataset. It needs more detail information to improve the chatbot so that it can generate more efficient and accurate output.

B. *Automatized Medical Chatbot (Medibot)*

This research is done by [2]. This research is proposed an automated medical chatbot which helps patient to determine the disease from their symptoms after the conversation with chatbot. This chatbot is based on knowledge base. Hence, after the chatbot has collected sufficient data from patient after questioning, it will generate the answer based the knowledge based. It means that the system will generate a match disease to symptoms that was given by patient. The training algorithms which used in the experiment are SVM,

KNN, Naïve Bayes. From the result, the SVM classifier get the highest accuracy which is 94.67.

Advantage: The system can help the patients to determine and recognize their disease so that they can prepare earlier and consult the doctor before it is too late.

Disadvantage: The system is a closed-source engine so when the new disease is not in the database the system will not recognize it.

C. BANK CHAT BOT–An Intelligent Assistant System Using NLP and Machine Learning.

[3] has proposed a 'BANK CHAT BOT - An Intelligent Assistant System Using NLP and Machine Learning' which uses NLP to develop a chat bot for banks. Users can send their query through the interface and the system will use NLTK library to run the pre-processing. NLP also uses tokenization to extract lemmas for each token. Vectorization will be proceeded to convert the text-format query to vectorized format. Then, ML logic is used to classify the query which class it belongs to. Based on the similarity values of the query, the matched answer will be reponsed to the user. The second feature of the system is that it has a feedback system to handle the query that is not able to be answered by the system. Thus developers can carry on to improve the system. Authors have compared over 7 algorithms to find out the best classification to improve the response efficiency. Grid search and randomized search are used for optimizing the algorithm. The proposed system has combined Natural language processing, vectorization and classification algorithms to reduce the processing time.

Advantages : Feedback system can help the developers to understand which part of the system needs to be improved. The combination of three algorithms make the proposed system be more efficient and do not require to train model every time.

Disadvantages: Not able to respond with images, links.

D. Implementation of a Chat Bot System using AI and NLP.

[4] has proposed a chatbot system for college inquiry system in paper " Implementation of a Chat Bot system using AI and NLP"(Lalwani, 2018). The chatbot has few features included being able to answer college admission related queries,profile viewing, retrieve attendance, provide information about examination to students, and fetch particulars about placement activities. The proposed system has three module which are Personal Query Response System, AIML Response System, and Query Analysis and Response System. The first module is to let the user get the specific information after he or she has pass the user authenticates successfully. AIML response systems can make normal conversation with the user based on the AIML file. AIML file is a knowledge base that is created to store the answer and question. Third module is check the similarity of the sentence with the predefined question. If the confidence is larger than 0.5, then the predefined answer will be send as a response.

Advantages: The system saves time for the student and staff in solving certain problems. For example, course details checking is able to be done easily by using the system.

Disadvantage: If the user input is not found in the predefined question set, the system might not answer properly. The system is not able to learn and save the new question that not in the AIML file.

E. Home automation using IoT and a chatbot using natural language processing

In the research of Home automation using IoT and a chatbot using natural language processing, [5] have proposed using a chatbot algorithm for users to text information to control the functioning of electrical appliances at home. The implementations of the chatbot are divided into two phases which are setup phase and use phase. In the setup phase, the questions given to the users will require users to provide information needed by the chatbot such as number of rooms or the number of appliances. Then in the use phase, various Natural Language processing techniques are applied to the text given by the user. They are tokenization, conversion into lowercase, removal of stop-words, Defining keyword and action lists, understanding keyword and action and actuation. Hence when the user input command, the web client and the application for the chabot will interact with the web server. The queries were received, processed on the server and sent to Raspberry Pi to perform control tasks. The Raspberry Pi then sends the acknowledgement signal back to the server and client when the task is completed. The user will then be notified about the successful completion of the task.

Advantage: This system is easy to be implemented into smart home appliances and brings convenience to those disabled people as it enables them to control the appliances without the need to move around frequently.

Disadvantage: only involved limited smart appliances such as lights, fan, television and air conditioning.

F. The Combination of Natural Language Processing and Entity Extraction for Academic Chatbot

[6] have proposed using an android-based chatbot to simplify the process of accessing academic information for students in the research of The Combination of Natural Language Processing (NLP) and Entity Extraction for Academic Chatbot. In this research, Natural Language Processing algorithm is applied as it is able to tag the parts of the conversation that contain verbs or nouns and look for the entities needed. NLP uses entity extraction to describe ambiguity in the human language. Entity extraction is an information recognition technique that refers to the process of identifying and classifying key elements of text into

predetermined classes and hence helps the machine to read the texts. In that research, the application experiment used a black-box testing model that focuses on the functional requirements of the created application and involved five students with different intake, faculty and major. The chatbot was able to answer most of the questions asked and only some were unrecognized due to the knowledge base system not up to date.

Advantage: The utilization of NLP in chatbot makes chatbot can accept and respond to inputs that are not exactly the same as the one in the knowledge base and can tolerate typing errors.

Disadvantage: Only specific and limited questions able to be answered by the chatbot.

G. Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners.

A research journal computed by several students from multiple universities mainly by [7] conducted an experiment about comparison between chatbot and human task partners. The aim of the experiment is to study whether chatbots are an effective learning tool to help students in education. Students are divided into groups with either chat bots or humans to complete a speaking task and provide their feedback from their learning efficiency and interest towards the subject. A structural equation model is constructed to keep records from the experiment such as score factors from students and scale relations. Since the experiment is conducted in 3 tests, the results are combined from the 3 tests for this experiment. The results are completed by using anova method and it shows that the interest score has no big difference until t3, which means the calculation should be considered using the mean difference to further expose the difference which results in human interactions are slightly better than chatbots but close to zero difference.

Advantages: Very convenient to use and assist on minor problems that doesn't require human resources to be involved.

Disadvantages: Chatbots are zero feelings bots that are unlikely to increase interest in the study field or even replace the education system directly which is close to useless at this rate.

H. A medical chatbot.

[8] had conducted a medical chatbot to solve users that submit their problem about their health situation by using text or vocal techniques such as natural language processing. The steps of using the systems start from the user login interface to register for the chatbot application. Next, users can ask questions about their health which is voice- text related. The systems complete the questions by running SVM algorithms and conclude the best outcome which is the most suitable medicine for the user. The purpose of this chatbot is not only for users with health issues or doctors to analyze but also for people who are interested to learn something about their health since it can be operated anytime. So by using SVM algorithms to predict disease and NLP to help communicate with humans made the systems technically workable and the accuracy is above 80%.

Advantages: Convenient on consulting health issues instead of users needed to be directly in front of doctors to consult health issues.

Disadvantages: Certain specific cases might wrongly lead to users thinking that they might have some critical disease which could be very costly.

III. DATASETS

This project will use one dataset to run and train the model. This dataset is taken from towards data science website (https://towardsdatascience.com/how-to-create-a-chatbot-with-python-deep-learning-in-less-than-an-hour-56a063bdfc44). It is stored in a form for json file which called intents.json. There are a series of tag which correspond to each own patterns and responses. In this case, the dataset is used to train a chatbot that used for pharmacy environment. Hence, the contents in the dataset are about the conversation which normally will use in the pharmacy. The below figure is a part of the dataset which showed the structure of it.

Figure 1: Example contents of tag patterns and responses

IV. SYSTEM FUNCTIONALITIES

A. *Text pre-processing (Lemmatization / Stop words removing)*

Before training, the data text needs to be clean because there may exist noise in the data. The experiment has provided few pre-processing steps to clean the text. Firstly, the text will be lemmatized so that the text will be neat and easy to process during the training. Lemmatization means that the words will be changed to the original words. For example, the "collected" will become "collect" after lemmatization processing. Then, the stop words removing step also will be conducted in the experiment. Stop words mean the words such as, a, an, the and etc. It is because the stop words are considered as the meaningless words in the text.

1) Tokenization

Tokenization means the sentences will be separated into tokens and each word become represent a token and stored as corpus for training. In the experiment, NLTK tools will be applied to process this step. The package in NLTK has its own default settings.

2) Train test split

In the experiment, the data has been split to 80:20. 80% for training and 20% for testing. After that, the data will be fed into the model for training.

B. *Model training（Two models）*

1) Model 1: Multilayer Perceptron (MLP)

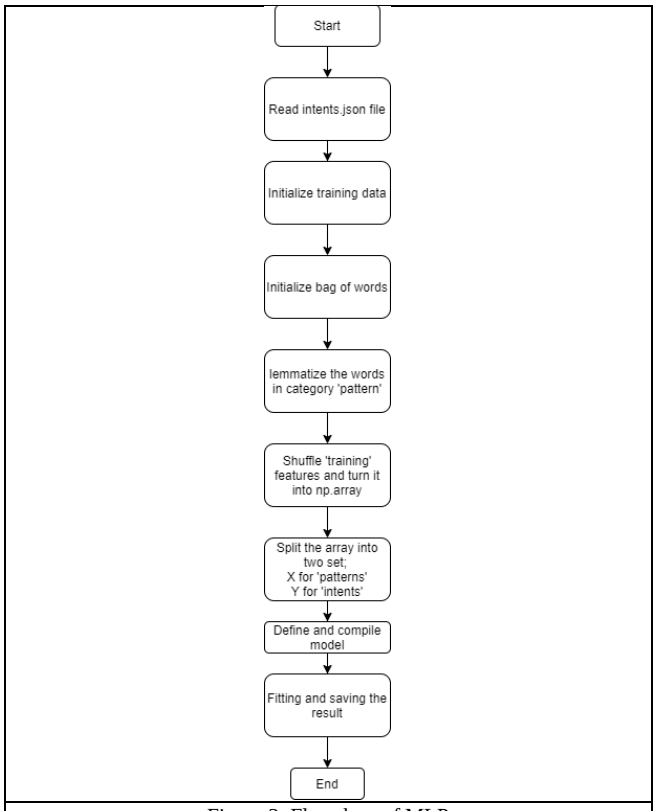

Figure 2: Flowchart of MLP

Multilayer Perceptron is a feedforward artificial neuron network and also a supervised learning method. The structure contains input layer, hidden layer and output layer. The input layer is also called as visible layer which received the input value and passed it to the following layer for calculation. The hidden layer is hidden inside the network. It can be one or many layers. The number of the layer will affect the performance of the model. More layer in the model will cause the training time increase. The output layer will receive the value from hidden layer and then classify the data. These are the brief concept of the MLP model.

This project will use a simple MLP method to train a model for the chatbot. The MLP contains three layer and each layer has different neuron. The first layer has 128 neurons, the second layer has 64 neurons and the last layer has 21 neurons. The number of the last layer neurons will follow the number of the classes label. In this case, the data has 21 labels so it shows 21 neurons. Due to the multilabel classification, the activation function of the last layer is set to "softmax".

To reduce the overfitting problem, the dropout layers have been added between the layer so that the model will not be complicated. The probability of the dropout layers has been set to 0.5. This model is using Stochastics Gradient Decent (SGD) as the optimizer. The learning rate is 0.01, decay rate is 1 e^{-6} and the momentum is 0.9. SGD optimizer is having a less memory requirement because it only takes one point at once during the derivative computation. The below equations show the equations for updating the weight (w_t) and bias(b_t) in SGD.

$$w_t = w_{t-1} - \eta \frac{\partial L}{\partial w_{t-1}}$$
$$b_t = b_{t-1} - \eta \frac{\partial L}{\partial b_{t-1}} \quad (1)$$

This model will use 100 epochs and batch size of 5 to train and test the model. Below figure are the structure of the MLP model.

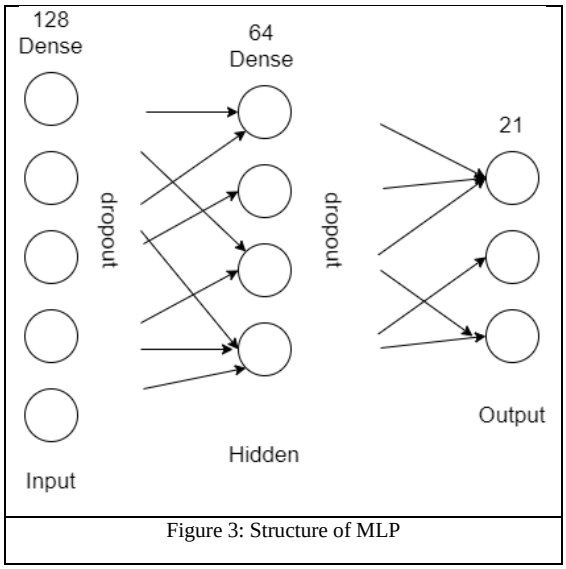

Figure 3: Structure of MLP

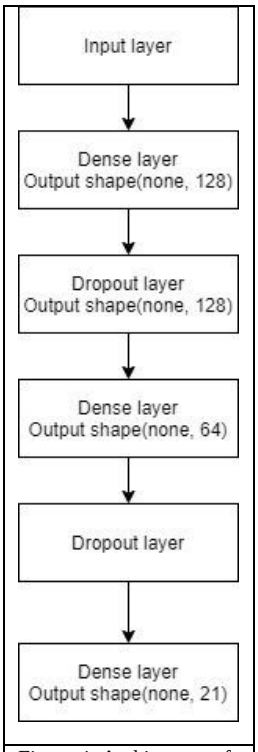

Figure 4: Architecture of MLP

2) Model 2: Convolutional Neural Network (CNN)

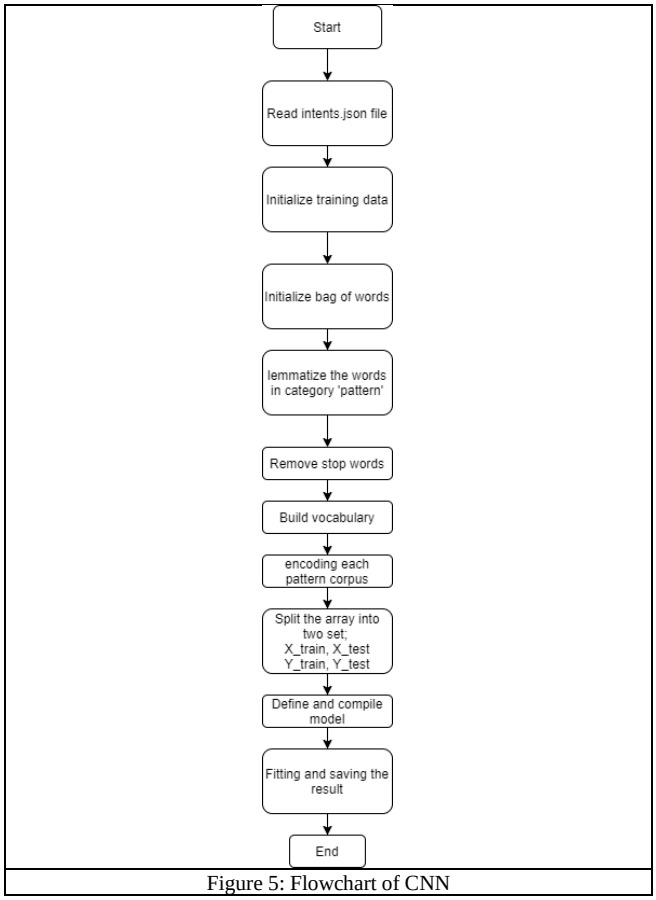

Figure 5: Flowchart of CNN

a) Embedding layer

The purpose of this layer is to create a word embedding layer which use a real numeric value in a high dimension to represent the words because computer is more deal with numeric data efficiently. The words which have similar meaning will get a closer vector value. In this model, the value of in each dimension will be random. There are 300 of dimension to represent a word and the maximum value of each embedding will be 16. The sentences in the dataset are short so 16 length is enough for it. The below figure shows the example of the embedding layer.

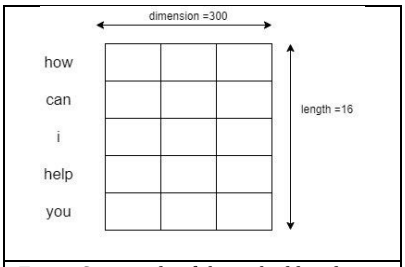

Figure 6: example of the embedding layer

b) Convolutional layer

The input of convolutional layer is the embedding layer. This layer is main layer the model. It is used to extract the important feature and produced the feature map. This model will use 64 filter following with 4 value of kernel size to deal with the data. The activation function used in this layer is Rectified Linear Unit (ReLU). After the feature maps have been produced, it will be passed to the max pooling layer.

c) Max Pooling layer

This layer will use the filter of 8 x 8 to scan through the feature map. The purpose of this layer is to reduce the computation burden by selecting the maximum value from the feature map to proceed to the next stage.

d) Flatten layer

This layer will take responsibility for changing the 3D tensor to 2D tensor. From the structure below shows that the output shape in the max pooling layer is 3-dimension. After going through the flattening process, the output shape becomes 2-dimension.

e) Dense layer

This layer is also known as a fully connected layer. In this model, there is only one layer that will be used with 22 neurons. To solve the multilabel classification, the activation function will use SoftMax function.

f) Other parameters

To decrease the overfitting problem, the early stopping with patience of 10. Hence, the model will stop training while the validation loss stops improving. The “adam” optimizer with a learning rate of 0.01 will be used in this model. Adam optimizer is good at dealing with noisy data or sparse gradients because it use the moving average of the gradient to prevent the stuck local minima. Below equations shows the update weight and bias formula:

$$w_t = w_{t-1} - \frac{\eta}{\sqrt{S_{dw_t} - \varepsilon}} * V_{dw_t}$$

$$b_t = b_{t-1} - \frac{\eta}{\sqrt{S_{db_t} - \varepsilon}} * V_{db_t} \quad (2)$$

There are 100 epochs to train the model. To classify the multilabel classification, the categorical cross entropy has been used as the loss function.

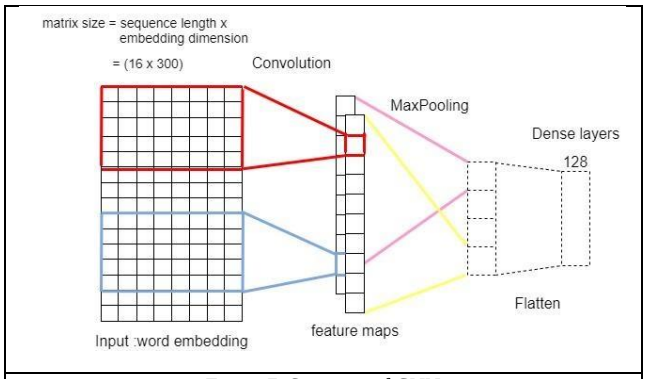

Figure 7: Structure of CNN

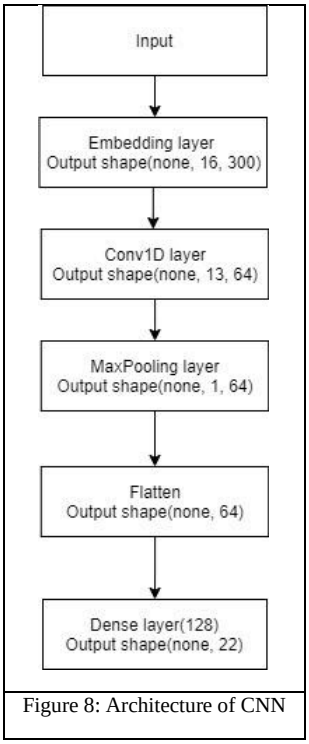

Figure 8: Architecture of CNN

C. *Creating Pharmacy Chatbot GUI*

After creating the model, the next step is to create an interface and apply the model on it. To create a GUI, a Python library, tkinter, is used and a chat window, scrollbar, send button, and textbox has been built by setting all the weight and height according to the design. Few functions are needed to make the GUI work properly. clean_up_sentence() function is used to clean the inputted sentence. bow() function takes the cleaned up sentences and creates a bag of words which apply for class prediction. The prediction is based on the model that was trained in the previous phase. In the function predict_class(), error threshold is set as 0.25 to avoid overfitting. predict_class() is used to generate a list of intents. the probabilities, and likelihood of matching the correct intent.Then, to generate responses, function getResponse() is created. This function will give the response with the highest probability through checking the json file. Lastly, to show the response GUI, chatbot_response() is created which includes predict_class() and getResponse() in this function. The sample of GUI shows as below:

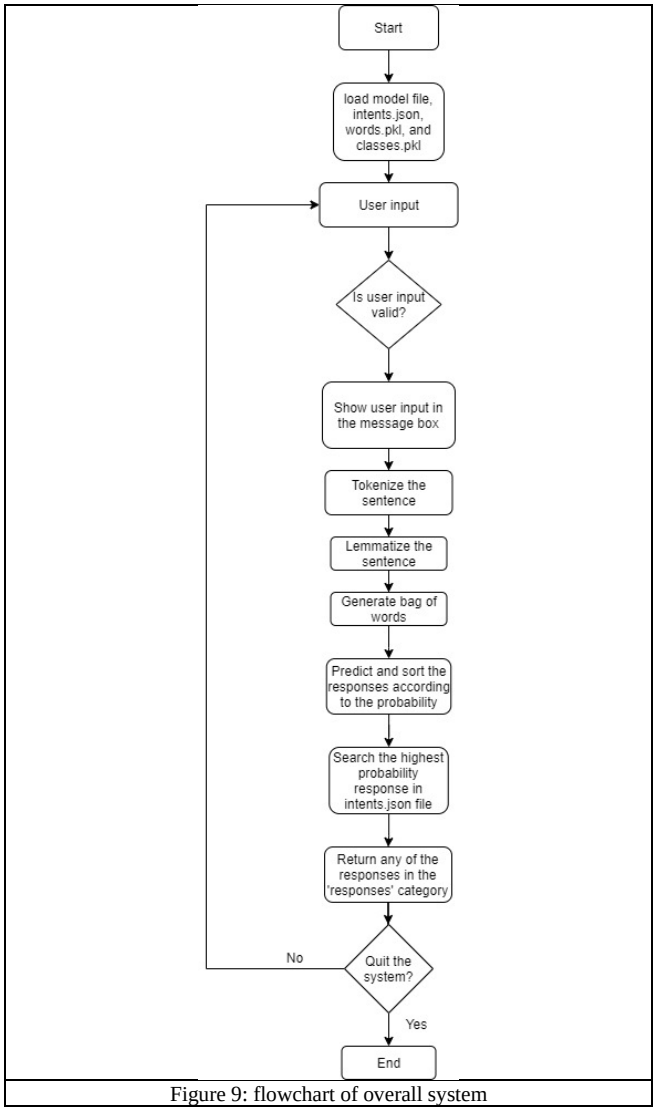

Figure 9: flowchart of overall system

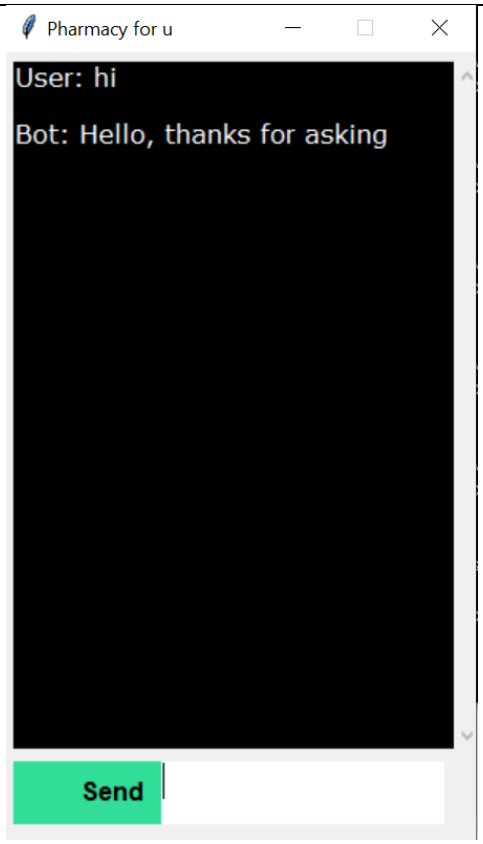

Figure 10: Chatbot GUI

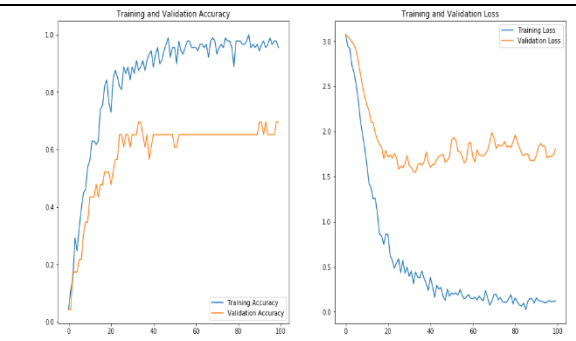

Figure 11: MLP loss plot

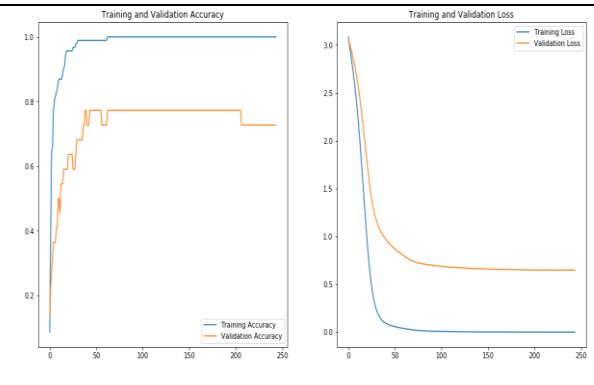

Figure 12: CNN loss plot

V. Results and Discussion

Table 1: Accuracy of trained models		
Models	MLP	CNN
Accuracy	69.57%	72.73%

As the table above shown the result of two model, CNN model get 72.73% accuracy which is higher than MLP model that get 69.57%. CNN model contains more high-level layer which can train the data deeper so it can perform better than MLP model. Besides that, the plot loss of each model shows MLP training and validation loss are unstable and having an overfitting problem whereas CNN model is more stable than MLP. This is because CNN model has applied the early stopping function which can prevent the overfitting problem efficiently. CNN also applied Adam optimizer which is more unstable at sharp minima than SGD optimizer.

References

[1] Patel, F., Thakore, R., Nandwani, I., & Bharti, S. K. (2019). Combating Depression in Students using an Intelligent ChatBot: A Cognitive Behavioral Therapy. *2019 IEEE 16th India Council International Conference (INDICON)*._J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Srivastava, P., & Singh, N. (2020). Automatized Medical Chatbot (Medibot). *2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and Its Control (PARC)*.

[3] Kulkarni, C. S., Bhavsar, A. U., Pingale, S. R., & Kumbhar, S. S. (2017). BANK CHAT BOT–An Intelligent Assistant System Using NLP and Machine Learning. *IRJET (International Research Journal of Engineering and Technology*, *4*(05).K. Elissa, “Title of paper if known,” unpublished.

[4] Lalwani, T., Bhalotia, S., Pal, A., Bisen, S., & Rathod, V. (2018). Implementation of a Chat Bot System using AI and NLP. International Journal of Innovative Research in Computer Science & Technology-IJIRCST, 6(3).Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[5] C. J. Baby, F. A. Khan and J. N. Swathi, "Home automation using IoT and a chatbot using natural language processing," *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, Vellore, 2017, pp. 1-6, doi: 10.1109/IPACT.2017.8245185.

[6] L. Tommy, C. Kirana and L. Riska, "The Combination of Natural Language Processing and Entity Extraction for Academic Chatbot," 2020 8th International Conference on Cyber and IT Service Management (CITSM), Pangkal Pinang, Indonesia, 2020, pp. 1-6, doi: 10.1109/CITSM50537.2020.9268851.

[7] Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, *75*, 461-468.

[8] Dharwadkar, R., & Deshpande, N. A. (2018). A medical chatbot. *Int J Comp Trends Technol*, *60*(1).