Using a Machine Learning Approach to Model a Chatbot for Ceylon Electricity Board Website

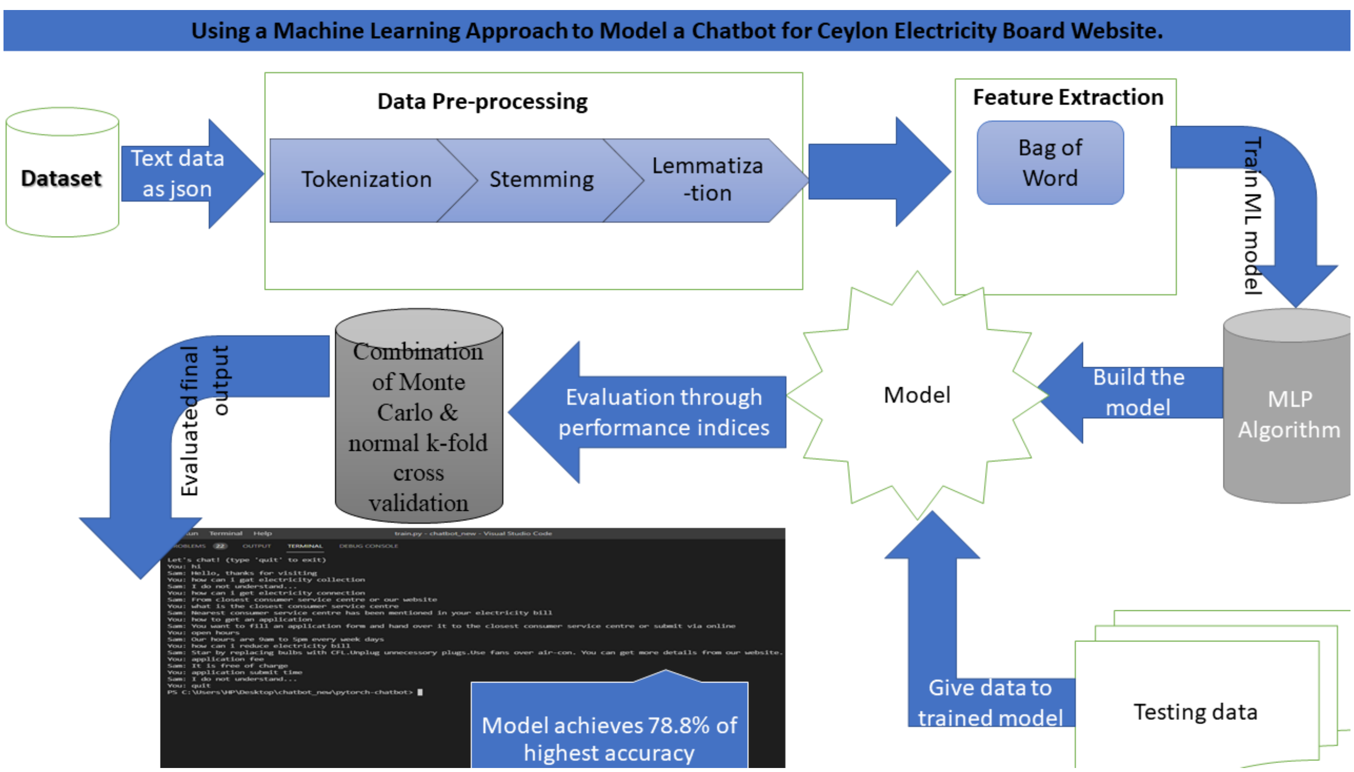
D.N.M. Hettiarachchi and D.D.A. Gamini

Department of Computer Science, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka

Corresponding author: gamini@sjp.ac.lk

Date Received: 16-08-2022 Date Accepted: 17-04-2023

Graphical Abstract



Abstract

Customer support is one of the main aspects of the user experience for online services. However, the rise of natural language processing techniques, the industry is looking at automated chatbot solutions to provide quality services to an ever-growing user base. In Sri Lanka, Ceylon Electricity Board website is one of the largest websites that customers use always to get information about electricity services. Hence, a chatbot system is very essential in CEB website. This paper presents a study about implementing and evaluating of a chatbot model for CEB website. This study implements virtual conversation agent based on deep learning algorithm which is multilayer perceptron neural network and a special text dataset for conversations about CEB services. The conversation agent model is made by utilizing the natural language processing techniques to facilitate the processing of user messages. The output of this research is the response from the chatbot and identify the best testing method to get highest accuracy for chatbot model. The chatbot model achieves the highest accuracy with the number of epochs set to 2000 and the learning rate value of 0.01 on response context data training so that it gets 78.8% accuracy.

Keywords: *Natural language processing, chatbot, deep learning, multilayer perceptron neural network, Monte Carlo cross validation, k-fold cross validation*

1. Introduction

Todays world technology plays a major role in all aspects of human life. Chatbot is the next big thing in the modern era of conversational services. Chatbot is also known as “chatterbot or conversational agent”. It is a software application used to conduct chat conversations via text or voice providing direct contact with human. Chatbot acts as an intelligent human but it makes harder for others to understand their real nature. Usage of chatbot has been widely expanded because of the development of more chatbots with various architectures and capabilities. In order to understand the user input and provide a meaningful response, chatbot uses artificial intelligence and deep learning methods. Moreover, they interact with humans using natural language (Ayanouz et al., 2020). Chatbots or conversational interfaces or digital assistants are form of intelligent virtual assistants which are taking over today's technology industry. The well-known Google Assistant, Apple's Siri, Amazon's Alexa and Microsoft's Cortana are the most famous voice driven virtual assistants. The text-based chatbot are also taking place over messaging apps and social media. Facebook Messenger released over 10,000 chatbots in less than a year. ‘Pandorabots’ which is a leading chatbot platform and community has created over 300,000 chatbots by more than 25,000 registered bot developers, as well as over six billion interactions with users have been recorded. Also, over 30,000 bots are reported as monthly activity conversational artificial intelligence agents in Microsoft. Approximately, 30 million messages are handled across thousands of company platforms, including UPS, Stack Overflow, Asiana Airlines, and many more (Schar, 2018).

Main usage of chatbots are dialog systems. For customer service, request routing and information gathering are various purposes for using chatbot as dialog systems. Further, some technologies and methodologies such as extensive world-classification processes, natural language processors and sophisticated AI are used by chatbots for simply scanning general keywords and generating responses using common phrases obtained from an associated library or database.

Most chatbots are accessed on-line via website popups or through virtual assistants. They can be classified into usage categories such as commerce, education, entertainment, finance, health, news and productivity. So, in this research, a chatbot model for Ceylon Electricity Board website will be developed.

The Ceylon Electricity Board (CEB) is the largest electricity company in Sri Lanka. With market share of nearly 100%, it controls all major functions of electricity generation, transmission, distribution and retailing in Sri Lanka ("Ceylon Electricity Board", 2022). Since CEB provides more services for customers, customers always need to contact them via phone or by visiting CEB branches to overcome their problems. But customers face difficulties when trying to contact CEB officers. This research intends to apply the deep learning methods to train a chatbot for website of CEB. The customers can interact with the chatbot just like they do with another human and through a series of queries.

One of the interesting questions that is aimed to study in this research is how well the standard techniques of deep learning can be applied to train the chatbot system. The main goal of this research is to eliminate the traffic of customers contacting the CEB officers by developing human-machine communication system. Further, this research tries to contribute to the development of chatbot model in presenting information about electricity services which were previously presented in the form of a FAQs system to be more interactive and practical.

Almost all websites use chatbot systems as customer service supporters or as information delivers or as connectors between customers and their finances (Dass, 2018). The latest statistics of using chatbot all over the world shows; "According to Drift's 2020 State of Conversation Marketing report, usage of chatbots as a brand communication channel increased by a whopping 92% since 2019. 24.9% of buyers used chatbots to communicate with business in 2020, up from 13% the year before" ("25 Top Chatbot Statistics", 2021), and "Chatbot growth has been prominent across a number of industries, to the point where 1.4 billion people now use them on fairly regular basis" ("The Future Is Now – 37 Fascinating Chatbot Statistics", 2022). But in Sri Lanka, there is no conversational chatbot system in Ceylon Electricity Board website which always customers need to get update about electricity services. So, it is important to model a chatbot system for CEB website for customers.

The main research topics in natural language processing are user intention identification and information extraction. Recently with the help of artificial intelligence and deep learning especially deep neural networks, many researchers have developed a lot of self-learning chatbots. Deep learning concepts such as Deep Neural Networks (DNN), Recurrent Neural Networks (RNN), and Convolutional Neural Networks (CNN) have been used to treat the chatbot model problems commonly (Ayanouz et al.,2020).

The Multilayer Perceptron (MLP) is a fully connected class of feedforward Artificial Neural Network (ANN). MLPs are useful in research for their ability to solve problems stochastically, which often allows approximate solutions for extremely complex problems. Lixa, a retrievable-based chatbot model trained using MLP (Aquil et al., 2021) could understand and guide on course materials and assignment for students, and provide faster responses compared to a generative based model. The MLP is combined with a special text dataset to help it conversed with users smoothly. NLP techniques are used for data preprocessing while feature extraction is done using a bag of words. A chatbot with a retrieval-based model using multilayer perceptron neural network could respond to the questions according to the stored response pattern (Aquil et al., 2021). The chatbot system achieved the highest accuracy with number of epochs set to 200 and learning rate of 0.02 on response context data training so that it got 96.43% accuracy.

Long short-term memory (LSTM) and Gated Recurrent Unit (GRU) are deep learning algorithms that are used to model and implement customer support chatbots. Customer service is an important point in the user experience, especially on the online scene where most users are extremely demanding both in terms of response time and quality of the answers given. A scalable and easily maintainable multilingual chatbot could interface with a company's existing customer support software (Peters, 2020). Users receive free-form text messages to help them with the most common issues they encounter regarding the company products. Single LSTM, single GRU, inverted GRU input and two GRU layers deep learning algorithms are used to train the model after preprocessing of data. Results show that, single GRU layer with inverted input sequence as the intent classifier for the final architecture provided the greatest accuracy and trained faster than its LSTM counterpart (Peters, 2020).

It is found that stereotype content model (SCM) can be applied to the domain of chatbots (Schar, 2018). Two different chatbots are used to check whether high warmth and competence scores have a positive influence on trustworthiness of a chatbot. A web-based experiment and a questionnaire is developed to determine whether the SCM theory can be applied to the context of chatbots. The questionnaire is first sent to a group of six people in order to receive feedback on it. Some questions are changed and inconsistencies removed after receipt of their feedback. By sharing final version of the questionnaire to family, friends and students, and creating a sample, applying the SCM to the domain of chatbots is claimed. As a result, research gap in this field is closed and human-computer communication by displaying highly warm and competent chatbots is improved. Chatbots are not seen as more trustworthy, and hence, future research would have to focus on the trust aspects of chatbots (Schar, 2018).

According to referred literature, there is a gap for the chatbot model evaluation. Most chatbot models have been evaluated from users' responses. But it is better to have proper evaluation method using evaluation matrices. This research is being done to improve accuracy and make conversation between chatbot and user as close to real world conversations. The novelty of this research is the method of validation. Here, combination of Monte Carlo and normal k-fold cross validation method is used to validate the chatbot model.

2. Methodology

After understanding the practical problem, a comprehensive study was done about the problem background. Online articles and videos about some technologies and methods such as NLP techniques, feature extraction and machine learning algorithms like neural networks were referred to get basic knowledge. The methodology is stated in Figure 1 work flow.

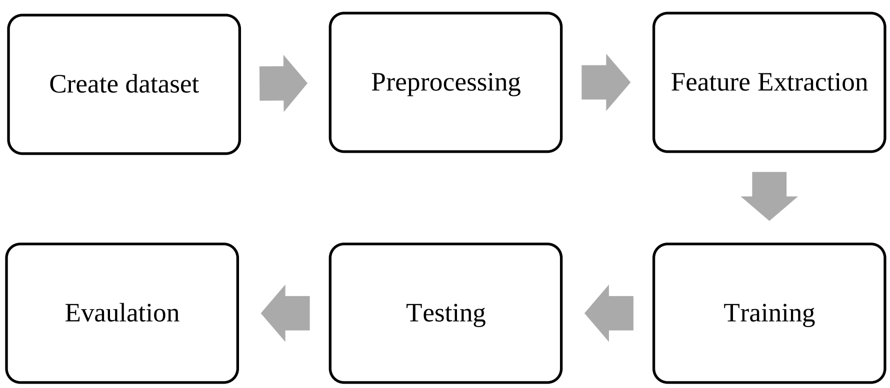


Figure 1: Work flow of the implementation

2.1. Create Dataset

Information about electricity industry and their services were gathered from CEB Website. While searching, "Supply Service Report of CEB" ("Home", 2022) which was published in CEB website was found. There was all information about electricity services, what customers should do when they have raised problems regarding power cuts, get new connections and bill issues. More details about CEB services and their procedure were explained in detail. Then, some questions and answers were listed by referring all resources that have been followed. Here, apart from the report, Frequently Ask Questions (FAQs) which were in CEB website were referred to create the dataset.

Basically, this dataset is a text dataset which contains words, phrases and sentences. Since text data falls under unstructured data category, there is no inherent structure. So, wide variety of words can be varied across the dataset and each sentence will also be of variable length as compared to a fixed number of data dimensions in structured datasets (Sarkar, 2019).

The chatbot dataset is designed using 250 patterns and some patterns have more than one response. Other than patterns and responses, it contains tags called as “class” that is used as a label by the model of search for related responses. Patterns are the questions that the user has to ask related to the existing class. Responses which are predefined for each class are the answers from the chatbot to the user. Since JSON or JavaScript Object Notation is a standard way of represent text-based data and it is easy to process, chatbot dataset has been stored as JSON file format.

2.2. Data Preprocessing

The inherent unstructured data (unformatted and noisy data) makes it harder for machine learning methods to directly work on raw text data. Hence, processing of data before training the model is very important, when it comes to the text data. Here, natural language processing pipeline have been used to data preprocessing or wrangling to remove unnecessary characters, symbols and tokens.

Natural Language Processing:

Natural language is the language human use to communicate with one another. On the other hand, human can tell machines what to do in a way machine can understand by using programming language. Natural language processing facilitates human to machine communication without humans needing to speak any other programming language as it allows machines to obtain and process information from written or verbal user inputs (“Natural Language Processing Chatbot Explained”, 2021). Furthermore, natural language processing is a field of Artificial Intelligence that helps computers understand, interpret, manipulate and respond to human in their natural language. Out of many natural language processing techniques, tokenization, stemming and lemmatization techniques were used to preprocess the chatbot dataset.

1. Tokenization: Tokenization is the process of breaking complex data onto simple units called tokens. A token can be a word or a letter or a number or a punctuation mark or a symbol. Mainly there are

two types of tokenization methods used in NLP, they are sentence tokenization and word tokenization. Here, the word tokenization which splits a sentence into list of words has been used (Chavan, 2020).

2. Stemming: Stemming is a normalization technique where list of tokenized words is converted into shorten root words to remove redundancy. That means in stemming stage, a token become stem by cutting prefix or suffix of the token. In this process, word stem is in their base or root form. A stemmer is a computer program that stems word. As an example, if we take there are three different tokens such as 'watching', 'watched', 'watcher' then stemmer reduces these words into the stem 'watch' (Chavan, 2020).

3. Lemmatization: Stemming always produces intermediate representation of word. That means stemmed word may or may not be meaningful. So, to overcome this problem lemmatization has been used. Here, lemmatization considers morphological analysis of the words and returns meaningful word in proper form (Chavan, 2020).

Data Cleaning:

Out of many available data cleaning techniques, textual data cleaning methods have been used to avoid noisy and inaccurate data. Removing tags and accented characters have been done as first cleaning step. It may contain unnecessary content like HTML tags, which do not add much value when analyzing text data. In English language, we have to deal with accented characters or letters like "é". So, by removing accented characters, these characters can be converted into standardized ASCII characters. Case folding process is another data cleaning step which is done to lowercase all letters and remove punctuation marks. Non alphanumeric character often adds to the extra noise in unstructured text data. So, it is preferred to remove them. In the English language, shortened version of words are known as contractions. Shortened words or phrases such as "don't", "I'll", "can't" etc. consider as noisy data in textual dataset. So, expanding contractions is an important step of data cleaning process. As a data cleaning step, removing whitespaces can also be taken to get a clear dataset. Finally, by removing stopwords, data cleaning process has been come to the end. Here stopwords mean less significance but most commonly use words in English language such as "a", "an", "the", "and" (Chavan, 2020).

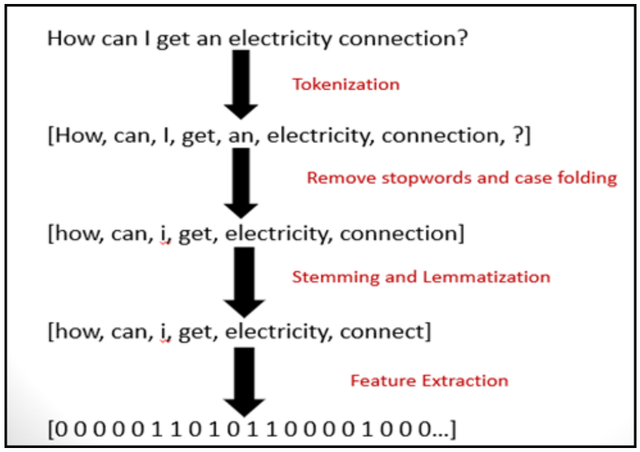


Figure 2: Preprocessing pipeline with an example

2.3. Feature Extraction

After the preprocessing phase, text data was needed to transform into meaningful vector of numbers. This is called as feature extraction of text data. In other words, it is the process of assigning a weight value to each word. Here, Bag of Words which is also known as a vector space model which is a simple representation used in natural language processing is used for feature extraction.

"The Bag of Words is a representation of text that describes the occurrence of words within a document" (Dass, 2018). Further, Bag of Words discard any information about the order or structure of words in the documents and only concern with whether the known words occur in the document. If the word in the list appears, it will be given 1 and if word is not appeared, then it will be 0 (Aquil et al., 2021). For example, let's take dictionary contains the words {electricity, is, not, use, good, thrifty} and we want to vectorize the text "use electricity thrifty", After model with Bag of Words, we would have (1, 0, 0, 1, 0, 1) vector.

The intuition behind the bag of words is that documents are similar if they have similar content and the meaning of the document can be learnt from its content alone (Dass, 2018).

2.4. Training the Model

After preprocessing and feature extraction, the result vector was processed in a classification method using a neural network with a feed-forward architecture, namely Multilayer Perceptron (MLP). Here, neural network with one input layer, two hidden layers and one output layer has been used. The input layer receives the input signal which is result vector of the feature extraction to be processed. Since this is classification problem model needs to classify which input vector belongs to which class. So, here outcome of the output layer is probabilities of the occurrence of input vector in each class or pattern. After comparing each probability, model selects the class of maximum probability. Hidden layers are placed in between the input and output layer are the true computational engine of the multilayer perceptron algorithm.

In fact, there is a theoretical finding (Lippmann,1987) that shows that an MLP with two hidden layers is sufficient for classifications.

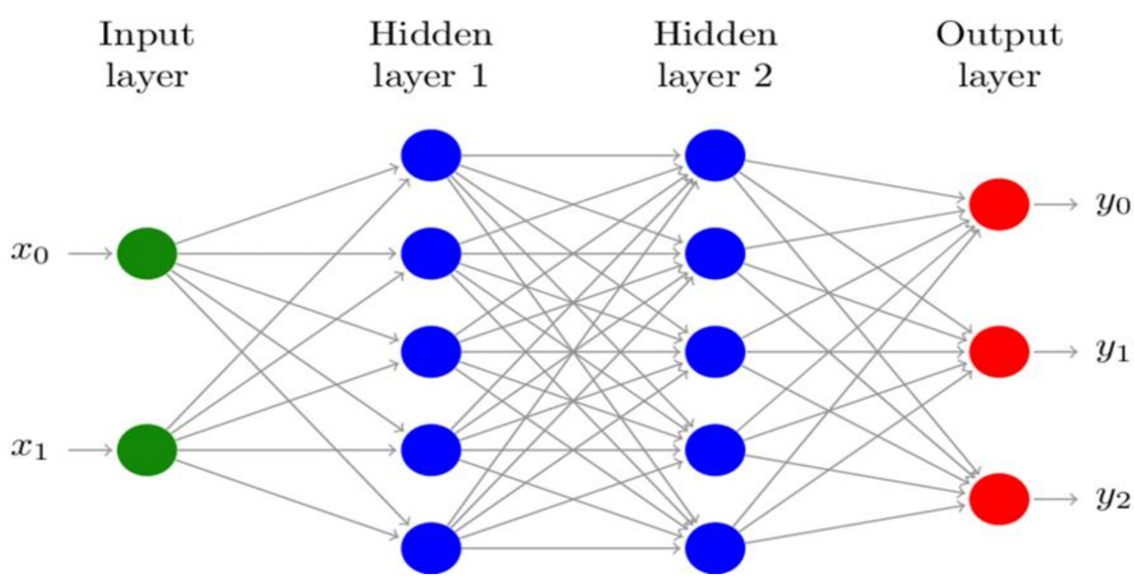


Figure 3: MLP architecture

Here, x_i: input vector and y_i: probability of the occurrences

Since multilayer perceptron uses a supervised learning technique, it is trained with the back-propagation learning algorithm. "Back-propagation is the practice of fine-tuning the weights of neural net based on the error rate or loss obtained in the previous epoch. Proper tuning of the weights ensures lower error rates, making the model reliable by increasing its generalization" (Al-Masri, 2019).

The chatbots are trained by storing question patterns according to responses and category. To move the information to forward, feed-forward architecture has been used in this model. The MLP model is mathematically interpreted as:

$$Output\ (O_i) = \sigma \left(\sum_{k=1}^{k} x_k w_{k,i} + \beta_i\right) \text{----------------------------(1)}$$

where, w: learning parameter, β: bias, x: input, k: number of input units, i: number of hidden units

ReLU non-linear activation function was used in the hidden layers and the output layer used softmax activation function to provide a probabilistic interpretation and generate multi-class outputs (Aquil et al., 2021).

2.5. Testing the Model

After training the model, chatbot model is able to give response for input message. As chatbot users, there should be some evaluation matrix to evaluate the accuracy of the responses in the chatbot model. So, chatbot model is needed to have testing methods in place to ensure accurate responses are given to users. In this study, two testing methods were used as evaluate the performance of the model. They are 80% / 20% split and combination of Monte Carlo and normal k-fold cross validation.

80% / 20% split:

The 80% / 20% split is the most basic approach which is 80% of data of the entire dataset involving of training process and 20% of data test the accuracy and ensure model an accurate chatbot. Accuracy of a chatbot is defined as the percentage of utterances having the correct intent returned.

Combination of Monte Carlo and k-fold validation:

1. K-fold cross validation: It divides the dataset into K number of parts, then uses one-fold at a time as the testing fold and the rest of the data as the training data. In this study, 5-fold validation has been used for testing the model. This means the dataset is split into 5 folds and the model is trained with 4 of the folds, and fifth fold is used to test. Repeat this until each fold has a turn as the testing fold. After averaging the overall accuracies of the folds together the accuracy of the chatbot is obtained.

2. Monte Carlo cross validation: It is similar to K-fold except the datasets are determined randomly. Here, the dataset is randomly split into an expected ration. That means, shuffling the data randomly and first 80% is selected as training data and rest 20% is picked as testing data. If we repeat the shuffling multiple times, we can have many splits as we want. Here, the dataset is shuffled 5 times and 5 accuracies are obtained to average together (Desmarais, 2018). In this research, to test the model and evaluate the result, Monte Carlo and normal K-fold test have been mixed together. By using Monte Carlo, the data is shuffled before the folds are created, but the folds are still unique because of the normal K-fold test.

3. Results

As mentioned in above section, 80% / 20% split test and combination of Monte Carlo and K-fold cross validation has been used to test the performance of the model. In this study, accuracy has been used as performance matrix of both testing methods. Since this is text-based classification, basically accuracy means the percentage of texts that were categorized with the correct tag. That is, the probability of containing user input and system response in same class or tag.

80% / 20% split:

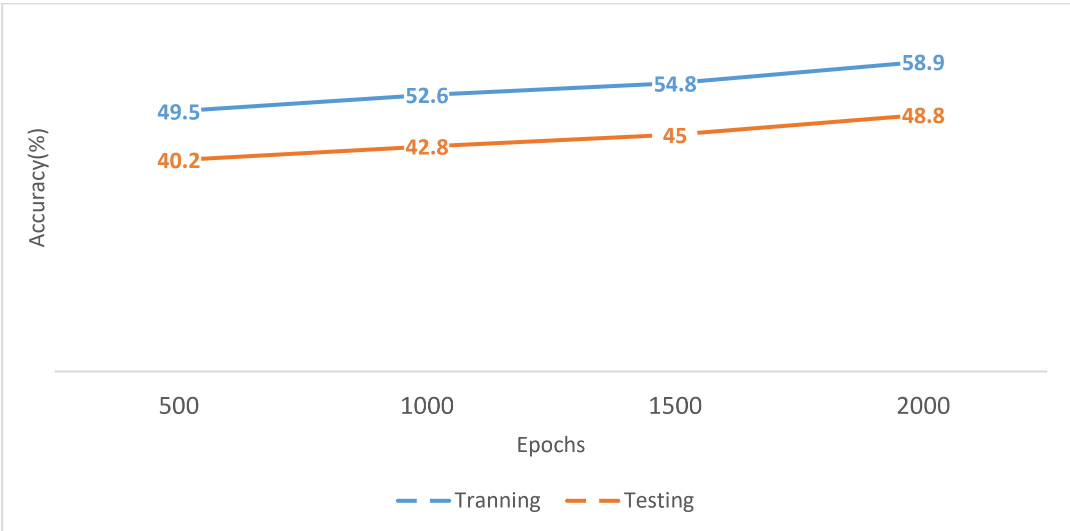


Figure 4: Accuracy of the model according to the number of epochs on 80% / 20% split

Here, we used learning rate as 0.01 and run the model for four epochs as mentioned in the graph in Figure 4. Since this test method is too simple and easy to implement, it gave 58.9% training accuracy and 48.96% testing accuracy in 2000 epochs. It could be seen from the same graph that the accuracy is gradually increasing with epochs.

Combination of Monte Carlo and normal k-fold cross validation:

Because 80% / 20% split testing method gave low accuracy, combination of Monte Carlo and normal k-fold cross validation method had to use. 0.01 of learning rate has been used to model the data and ensured accuracy for different epochs as before. In this testing method, 5 folds have been used with random selection.

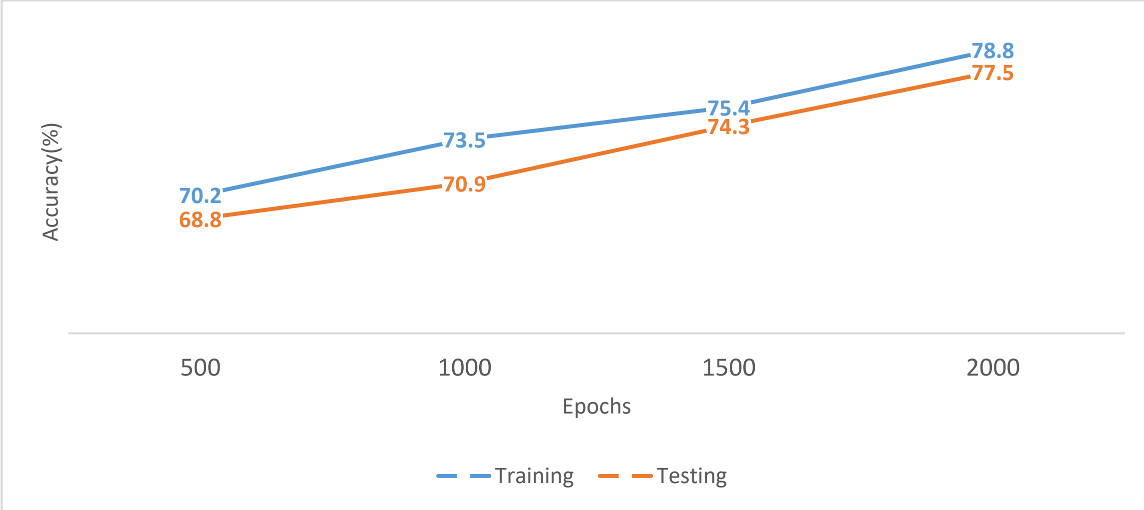


Figure 5: Accuracy of the model according to the number of epochs on Combination of Monte Carli k-fold cross validation

Practically, K-fold cross validation is bit more complicated to implement since we have to create each fold. But as shown in Figure 5, it gives more accurate results which is 78.8% of training accuracy and 77.49% of testing accuracy. As in 80% / 20% split, this testing method is also gradually increasing with respect to epochs.

The principal goal of machine learning is to create a model that performs well and gives accurate predictions in a particular set of case. In order to achieve that, machine learning optimization is needed. Machine learning optimization is the process of adjusting hyperparameters in order to minimize the cost function. Hyperparameters are set before starting to train the model. Out of many hyperparameters, learning rate and epochs of the model were tuned manually to improve the performance of the model.

Epochs are the number of stages required for the system to carry out the training process. Testing was started with 500 epochs and a leaning rate of 0.01. After that the number of epochs was chosen as 1000, 1500 and 2000. The result in Figure 6 shows that the accuracy of the model is increasing as the number of epochs increases. The highest level of accuracy, which is 78.8%, is obtained with the number of epochs 2000. This is because as the number of epochs increase, the chatbot system will learn to reduce the error value and learn to increase the accuracy value more.

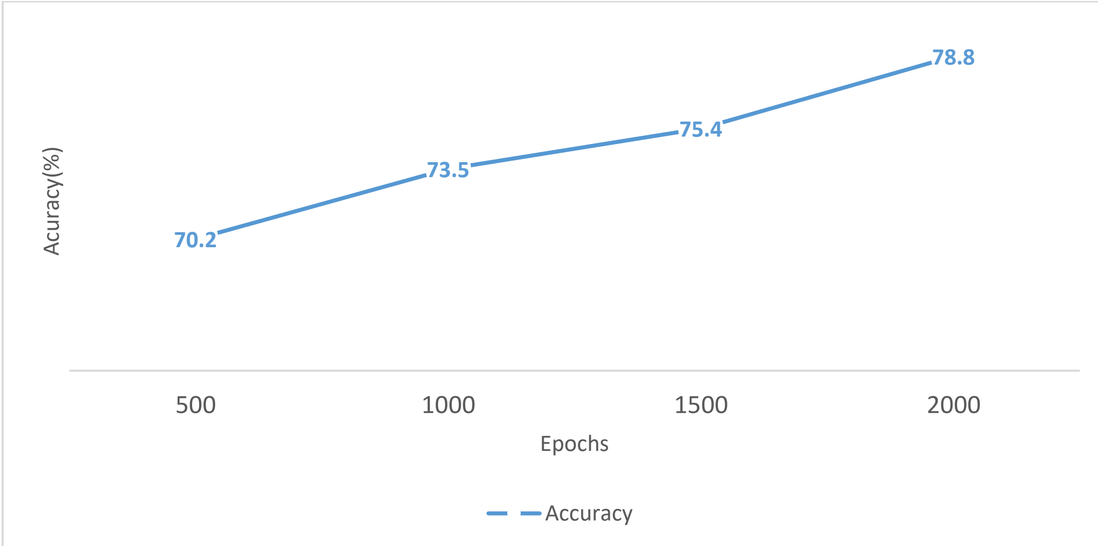


Figure 6: Accuracy of the model according to the number of epochs

Learning rate is the level of model accuracy to correct the error value or to minimize the loss value. In order to see how the accuracy varies, different learning rates were tried with the number of epochs as 2000 which gave the maximum accuracy value previously. The learning rates tested are 0.1, 0.02, 0.01 and 0.001.

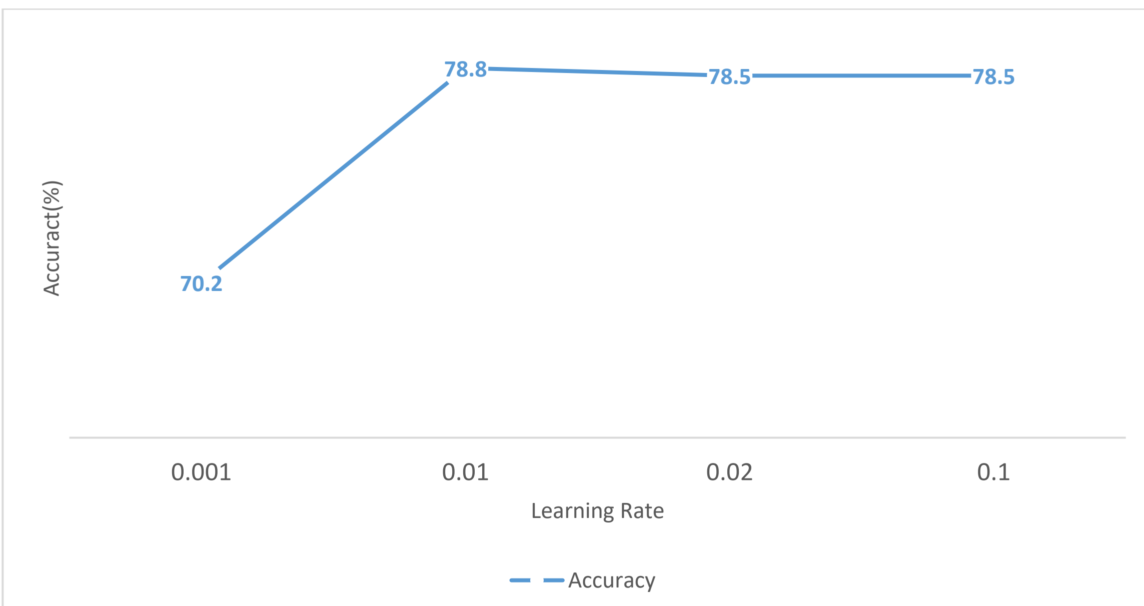


Figure 7: Accuracy of the model according to learning rate

As mentioned in the Figure 7, the highest accuracy levels which are 78.8%, 78.5% and 78.5% obtained with a learning rate value of 0.01, 0.02 and 0.1 respectively. This is because the smaller the value of learning rate, the greater the accuracy to reduce the error value in the system. But it will increase the training process time and require a larger number of epochs to reduce the loss rate to convergent so that the loss vaule decreases.

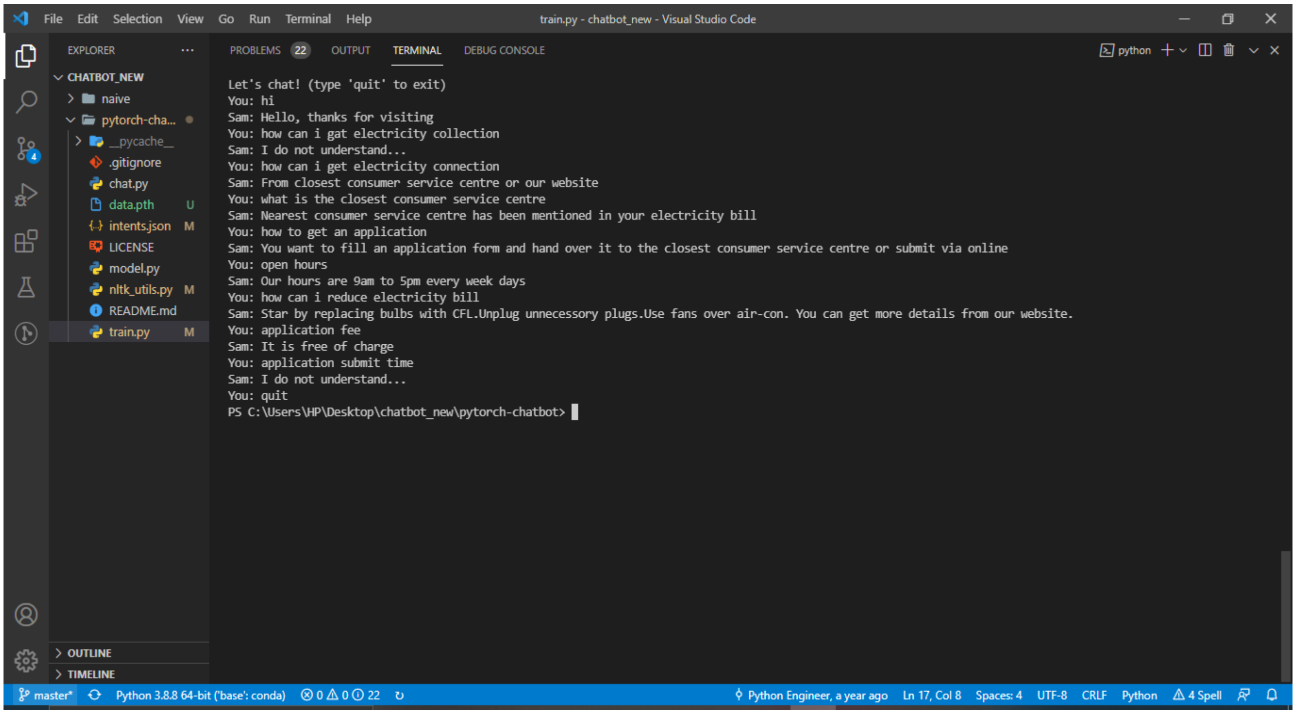


Figure 8: Chatbot model testing

As seen in Figure 8, there is a conversation that user had with the implemented chatbot model. Here, the modeled chatbot simply understands the user's questions and gives the particular answer after the knowledge it has learnt from the training.

4. Discussion

Traditional conversation services to customers often emphasize the need for excessive information on users, besides that traditional conversation to users often take a lot of time and cannot be served 24 hours. Basically, when customers need to get information about electricity problems, they always face a lot of difficulties because of traditional conversations. Therefore, a chatbot model was implemented to respond to several things regarding the services of Ceylon Electricity Board. This study examines the addition of machine learning to chatbots that are useful for conducting rapid evaluation and analysis of data.

According to the result of this study, combination of the Monte Carlo and normal k-fold cross validation is better than 80% / 20% split testing method. Accuracy is increased over 25% in combination of the

Monte Carlo and normal k-fold cross validation compared to 80% / 20% split. The model performed best accuracy of 78.8% with 2000 epochs and a learning rate of 0.01.

The main objective of this research is to model the neural network model based chatbot in providing appropriate responses. This chatbot retrieval model that can focus on answering only about the features of the CEB services.

5. Conclusion

This chatbot model can be used to interact with Ceylon Electricity Board Website users and it helps users to get appropriate responses related to their questions. This study found that a chatbot model using a multilayer perceptron neural network can respond to the questions according to the stored response pattern that has been created. The chatbot system can respond using a multilayer perceptron artificial neural network well in several text message tests with response pattern storage containing 250 different patterns. The novelty of this study is using combination of Monte Carlo and normal k-fold cross validation to test the model.

The chatbot model achieves the highest accuracy with the number of epochs set to be 2000 and the learning rate value of 0.01 on response context data training so that it gets 78.8% accuracy.

The chatbot model created in this study has the ability to answer faster and is more accurate in providing answers to the questions that have been provided. Besides that, this chatbot model also has drawbacks such as not being able to respond appropriately to questions other than those already provided in the dataset. The chatbot also has ambiguity in some of the words that have been provided due to incorrect word pairs.

Suggestions that can be used for further research, especially in research related to the development of chatbot applications is the need to enrich the dictionary of words. In addition, it is also necessary to enrich the existing patterns in the knowledge base with a combination of different patterns and sentence structures so as to increase the possibility of a better response.

References

1. Aqil, A.N., Dirgantara,B., Ahmad, U.A., Septiawan, R.R., Suherman, A.L.,2021.Robot Chat System (Chatbot) to Help Users ‘Homelab’ based in Deep Learning. Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 8, doi: 10.14569/IJACSA.2021.0120870.
2. Ayanouz,S., Anouar,A.B., Benhmed,M.,2020.A Smart Chatbot Architecture based NLP and Machine Learning for Health Care Assistance. doi: 10.1145/3386723.3387897.
3. Al-Masri,A.,2019.How Does Back-Propagation in Artificial Neural Networks Work? Medium. https://towardsdatascience.com/how-does-back-propagation-in-artificial-neural-networks-work-c7cad873ea7. Accessed 07 March 2022.
4. Bhagwat, Vyas, A., 2018. Deep Learning for Chatbots. Master of Science, San Jose State University. doi.org/10.31979/etd.9hrt-u93z.
5. Chavan, J.,2020. NLP: Tokenization, Stemming, Lemmatization, Bag of Words, TF-IDF, POS. Medium. https://medium.com/@jeevanchavan143/nlp-tokenization-stemming-lemmatization-bag-of-words-tf-idf-pos-7650f83c60be. Accessed 03 March 2022.
6. Ceylon Electricity Board. Wikipedia. https://en.wikipedia.org/w/index.php?title=Ceylon_Electricity_Board&oldid=1073800571. 4 Feb 2022.
7. Dass, R.,2018. Create your chatbot using Python NLTK. Medium. https://medium.com/@ritidass29/create-your-chatbot-using-python-nltk-88809fa621d1. Accessed 8 Oct 2021.
8. Dhyani, Manyu, Rajiv, K., 2021. An Intelligent Chatbot Using Deep Learning with Bidirectional RNN and Attention Model. Materials Today: Proceedings 34:817–24, doi.org/10.1016/j.matpr.2020.05.450.
9. Elcholiqi, A., Musdholifah, A., 2020.Chatbot in Bahasa Indonesia using NLP to Provide Banking Information.IJCCS Indones. J. Comput. Cybern. Syst., vol. 14, p. 91,doi: 10.22146/ijccs.41289.
10. First Sinhala Chatbot in action. 2022. https://www.researchgate.net/publication/230688513_First_Sinhala_Chatbot_in_action?enrichId=rgreq-4c7efb5e748596322a1e71e6afa563be.
11. Home. 2022.https://ceb.lk/.Kbuell, A sentiment-based chat bot. 2020. Presentica. https://www.presentica.com/doc/11296301/a-sentiment-based-chat-bot-pdf-document.Accessed 18 Feb 2022.

12. Kuksenok,Kit,Andriy M., 2019. Evaluation and Improvement of Chatbot Text Classification Data Quality Using Plausible Negative Examples. In Proceedings of the First Workshop on NLP for Conversational AI, 87–95. Florence, Italy: Association for Computational Linguistics. doi.org/10.18653/v1/W19-4110.
13. Lippmann,R., 1987. An Introduction to computing with neural nets. IEEE ASSP Magazine. https://ieeexplore.ieee.org/abstract/document/1165576.
14. Natural Language Processing Chatbot Explained | Landbot. Landbot.io. https://landbot.io/blog/natural-language-processing-chatbot.Accessed 09 Oct 2021.
15. Pardeshi, Siddhi, Suyasha, O., Pranali, S., Manasi, B., Anandkumar,B., 2020. A Survey on Different Algorithms Used in Chatbot. 07, no. 05 (2020): 7.Sarkar,D., 2019. Traditional Methods for Text Data. Medium. https://towardsdatascience.com/understanding-feature-engineering-part-3-traditional-methods-for-text-data-f6f7d70acd41. Accessed 02 Mar 2022.
16. Schär,A., Chatbots/Conversational Interfaces in the Context of the Stereotype Content Model (SCM). 80.
17. Sheikh, Salim, 2019. Artificial Intelligence based Chatbot for Human Resource using Deep Learning a dissertation submitted in partial fulfilment of the requirements for the award of the degree of master of technology in advance computing and data science. https://doi.org/10.13140/RG.2.2.25465. 52322.
18. Singh, Shrushty, Prashant,M., Gupta, Rashmi,M., Pratyusha,S., 2021. AI Powered Chatbot Using RASA.
19. The Future Is Now-37 Fascinating Chatbot Statistics. SmallBizGenius. https://www.smallbizgenius.net/by-the-numbers/chatbot-statistics/. Accessed 02 Mar 2022.
20. Tiedekunta, I., Natural language processing techniques in chatbot development: how does a chat- bot process language? 72.
21. Vijayaraghavan, V., Jack, Cooper, Rian, J., 2020. Algorithm Inspection for Chatbot Performance Evaluation. Procedia Computer Science 171 (2020):2267–74. doi.org/10.1016/j.procs.2020.04.245.
22. Top Chatbot Statistics: Usage, Demographics, Trends. 2021. Startup Bonsai. https://startupbonsai.com/chatbot-statistics/.