PAPER • OPEN ACCESS

Retrieval Based Chatbot on Tarumanagara University With Multilayer Perceptron

View the article online for updates and enhancements.

Retrieval Based Chatbot on Tarumanagara University With Multilayer Perceptron

Andrew Ciayandi[*], Viny Christanti Mawardi, Janson Hendryli

Informatics Engineering Department, Faculty of Information Technology,
Universitas Tarumanagara

[*]andrew.535160062@stu.untar.ac.id

Abstract. Natural Language Processing or NLP is a branch of Artificial Intelligence that focuses on processing language. One of the implementation of NLP is in the field of customer service in the form of chatbots. Tarumanagara University itself already has a customer service called Admission Tarumanagara. Admission can be contacted by telephone or sending an e-mail message. Making Chatbot will add ways to get information about Tarumanagara University. Chatbot is a computer program based on artificial intelligence that can simulate conversations between people. The system created is a Retrieval Based Chatbot system that can recognize user questions and answer user questions directly. Making a chatbot system will use a model with the Multilayer Perceptron method. The process for the system will go through two stages, namely the preprocessing stage and the introduction stage. The preprocessing stage consists of case lowering, tokenizing, and stemming. After that the model is made and the introduction phase of the trained model is carried out. Chatbot that has been made can be accessed using the LINE application via LINE ID @764vvcfm.

1. Introduction

Information is data that is processed into a form that is very useful for making decisions. Information is useful for decision makers because information reduces uncertainty in the data, because based on that information the managers can know the objective conditions[1]. Before making a decision, sufficient information is needed as a helper to make a decision. Chatbot is designed to alleviate the role of humans in answering user questions so there is no difficulty in finding the information needed. Chatbots are programs that mimic human conversation using Artificial Intelligence (AI). It is designed to be the ultimate virtual assistant, entertainment purpose, helping one to complete tasks ranging from answering questions, getting driving directions, turning up the thermostat in smart home, to playing one's favorite tunes etc.[2]

Chatbot can be implemented in areas where conversation is narrow and not for open conversations. The solution is done by creating a chatbot with Retrieval Based Models. The Retrieval Based Model will help chatbots in giving the appropriate answer to the user compared to the Generative Model which produces its own text and results in irrelevant answers[3]. The system will use the Multilayer Perceptron machine learning architecture in order to build the chatbot because this architecture is suitable for category classification so that it can answer user questions accurately[4].

Marline, Junaedy, and Hadman use Artificial Intelligence Markup Language (AIML) in order to make chatbot[5]. By using this method, chatbot can only answer based on whether chatbot has the same pattern of sentences in the databasse. This type of chatbot usually use button in order to direct the user.

Suryanto uses Recurent Neural Network (RNN) and Named Entity Recognition (NER) in order to make chatbot[6]. This type of chatbot is called Generative Chatbot which produces its own output. The downside is often times the output doesn't have the correct structure in grammar and vocabulary. Generative chatbot is usually used as a companion chatbot.

The aim of Retrieval Based Chatbot is to classify user's question and give answer based on the classification it obtains by using machine learning.

2. Methods

The methods section contain enough information to enable the reader to understand what has been done, and important question to which the section should provide brief answers.

2.1 Design

This system is designed using the Multilayer Perceptron method which is one of the methods used to do machine learning. This method is suitable for classifying based on the specified category. In the initial stage the system accepts input questions in the form of text. After that the system performs text normalization processes such as Case Lowering, tokenizing, and stemming. The chatbot system scheme can be seen in Figure 1.
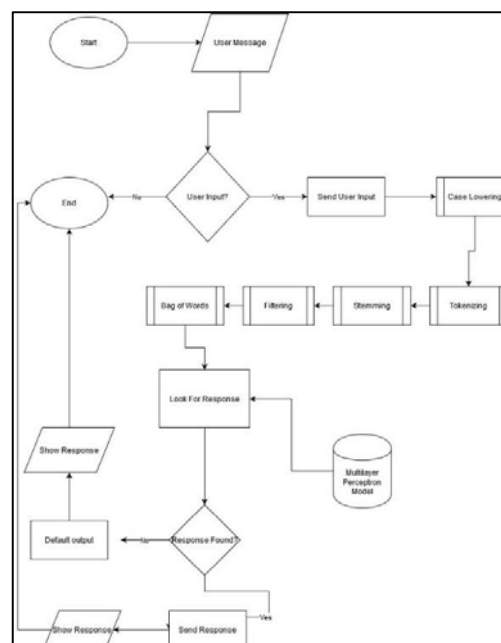

Figure 1 Chatbot System Scheme

2.2 Gathhering Data

The system will be built starting with data collection from the Tarumanagara University website with the link http://untar.ac.id/. After the data is collected, the data will be processed into JSON format. After that

the text processing will be done which includes text lowering, Tokenisation, and Stemming. Case lowering will make all uppercase letters into lowercase, tokenisation will convert sentences to words, and Stemming will make all words into basic and standard words. After doing the text processing, vocabulary will be made with the format of bag of words. After the BoW model will be formed, the BoW representation will be used to train using machine learning with the Multilayer Perceptron architecture. The scheme can be seen in figure 2.
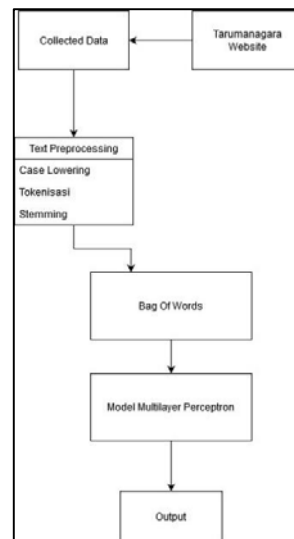

Figure 2 Data Scheme

2.3 LINE

The schema will explain the flow that will occur when the user sends a message. When a user sends a message, the message will be processed as input, and the input will be received by the LINE Messaging API. After that LINE will send the input to Heroku. Heroku will process the input data using a chatbot program that has been uploaded previously. After that the output of the chatbot program will be returned to LINE by Heroku. LINE receives output from Heroku which will then be forwarded to the user as a reply message. The scheme can be seen in Figure 3.To access chatbot, user must add chatbot as friend first. The ID of the chatbot is @764vvcfm.
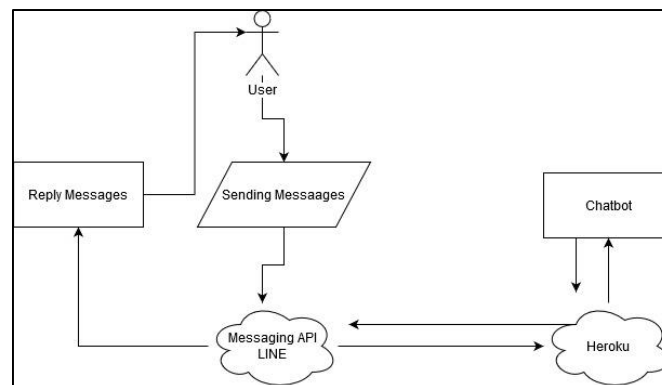

Figure 3 LINE Chat Scheme

2.4 Administration Website

Administration website is used in order to manage chatbot data. Hierarchy diagram is used to provide information about the basic structure of the website. Hierarchy diagram for the website can be seen on Figure 4.

Figure 4 Hierarchy Diagram

3. Results and Discussion

3.1 Website Testing

Website that has been made is used as a place to manage chatbot's data. In order to access the website, the administrator has to input his/her credential first in login page in order to make any changes in the data. At figure 5 is the interface for the Home and commit module. In this module, we can input new data into the chatbot which consists of tags, response and pattern.

Tags is used as an identifier, response is a example of question for the chatbot to train, and response is how the chatbot should responses when it encounters such question. For example, if the information we want to provide is about how to register as a new student in Tarumanagara University, we would first create a new tag "Registration", then we put question-like sentence into pattern textbox such as "How to register to Tarumanagara University", then we put the answer in response textbox such as "To register as a new student in Tarumanagara University, you could contact us at 62 21 569 58 723 or you could visit our website at admisi.tarumanagara.com", then we click "Simpan" button. In order for the chatbot to learn this new information, we first need to train it by pressing the "Create Model" button. After that we could press "commit" button to send the new chatbot to Heroku .
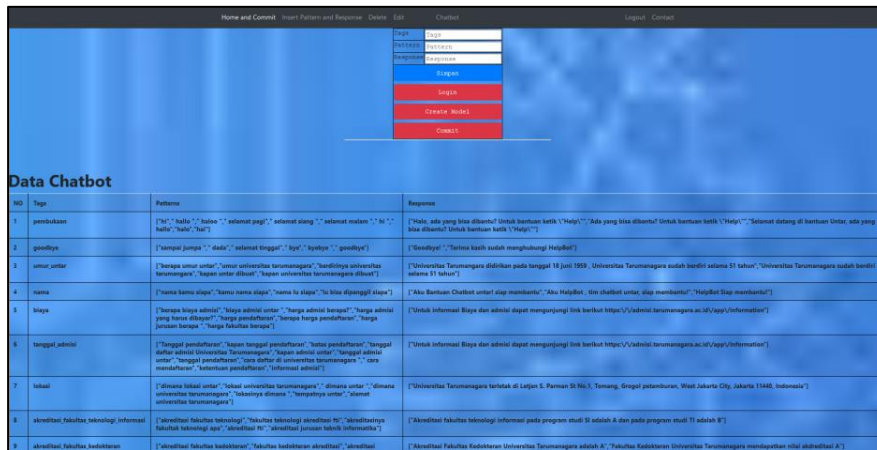
Figure 5 Home and Commit Module

There is also delete module in this website. In this module, administrator can delete past data. Administrator will use provided listbox which houses all tags, and choose whether to delete all data, only patterns or only responses on selected tag. At figure 6 is the edit module, which is used to make any changes on past data. Administrator will also use the provided listbox in order to choose which tag he/she wants to change. The administrator can choose whether to only edit the pattern by clicking "commit edit pattern" or edit the responses by clicking "commit edit responses".
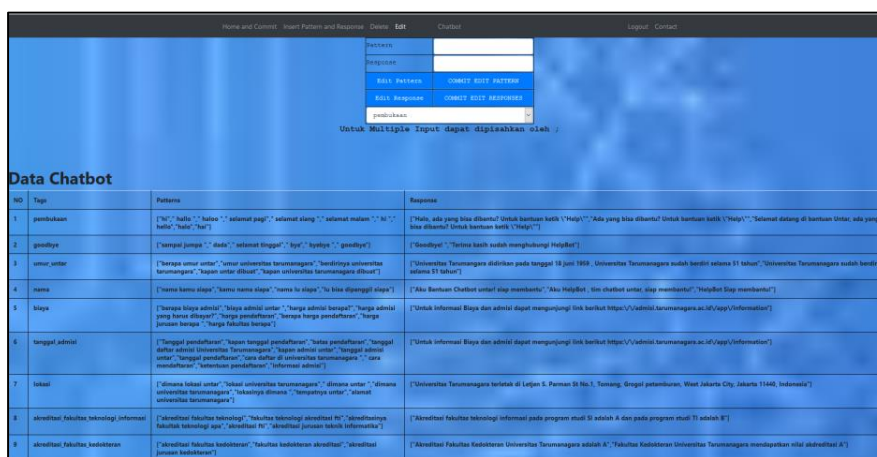

Figure 6 Edit Module

3.2 Line Chat Test

Line chat test is done with the aim of whether users can interact directly with Chatbot that has been created. In order for user to be able to interact with the chatbot, user needs to add the chatbot as a friend in LINE application with ID @764vvcfm. After that said user needs to open the chat window in order to chat with the chatbot. Figure 7 shows where users can give questions to chatbot.
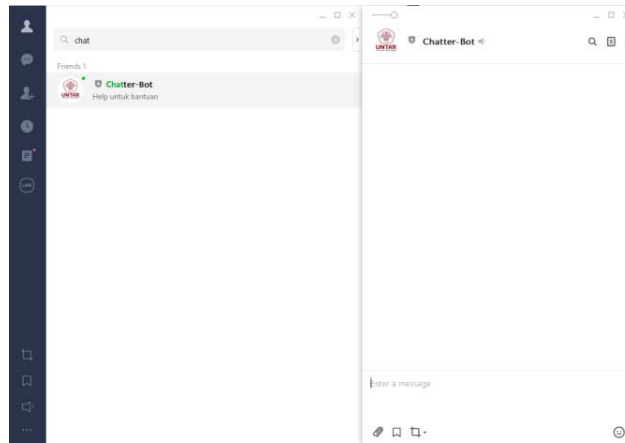
Figure 7 LINE

After giving it question, the chatbot will answer user's question immediately. "Help" can be given to chatbot in order to list the capabilities of the chatbot. If the chatbot has been trained on the particular subject, then the chatbot will answer it with the provided response, but if it is outside of the chatbot capabilites, then it will answer with "maaf saya tidak mengerti pertanyaan anda" sentences as shown in figure 8. There are a few limitations in the chatbot which user should follow such as:

1. User must use EYD Indonesian Language
2. Input must be longer than 1 word
3. Chatbot only provides information about Tarumanagara University

Figure 8 LINE Chatting

4. Conclusion

1. The result of this study is a chatbot that can provide answers to user's questions about Tarumanagara University. User can access the chatbot by adding LINE ID @764vvcfm.

2. In the LINE chat test, the chatbot has been made well and is able to provide a direct response to the questions given by the user.

3. The function of the administrator website that has been designed has been running properly. Administrator is able to add, change and delete data on the chatbot that has been created. Admin can also create new models using the admin website that has been created.

5. References

[1]   Transalation from Setiawan,Parta (2020). Pengertian informasi – ciri, jenis, sumber, kualitas,tujuan, manfaat, komponen, perubahan, contoh, para ahli.https://www.gurupendidikan.co.id/pengertian-informasi/, Accessed on February 2020, 25th

[2]   Ranoliya, Bhavika R., Nidhi Raghuwanshi, and Sanjay Singh. "Chatbot for university related FAQs." In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1525-1530. IEEE, 2017.

[3]   Britz, Denny. deep learning for chatbots, part 1 – introduction . http://www.wildml.com/2016/04/deep-learning-for-chatbots-part-1-introduction/, Accessed on February 2020, 25th .

[4]   Ramchoun , Hassan. new modeling of multilayer perceptron architecture optimization with regularization: an application to pattern classification . *IAENG international journal of computer science* 44(3) : 261-269.

[5]   Transalation from Wijaya M, Junaedy and Arfandy H. "Perancangan chatbot untuk informasi penerimaan mahasiswa baru pada stmik kharisma makassar", *Jurnal Ilmu Komputer Kharisma Tech*,Vol 14, Num 1, 2019.

[6]   Suryanto. "Perancangan Chatbot Dengan Reccurent Neural Network Dan Named Entitiy Recognition", Jakarta: Faculty of Information Technology, Tarumanagara University, Not Publicized, 2019.

Jakarta, 22nd July 2020

No.   : 027-TIM/2814/UNTAR/VI/2020

FULL PAPER ACCEPTANCE NOTIFICATION

Reference Number   : 162_TICATE 2020
Title   : Rertrieval Based Chatbot on Tarumanagara University with Multiplayer Perceptron
Authors   : Andrew Ciayandi, Viny Christanti Mawardi, Janson Hendryli

Dear Mr./Madam

Thank you for your paper submission to the TICATE 2020. We are pleased to inform you that your submission is **accepted** for publication and presentation in TICATE 2020. Due to some outstanding reviews for other submissions, however, the review result is not yet available. Currently the review is still on progress. In the meantime, we recommend that you check your manuscript to minimize obvious errors, such as formatting and grammatical errors.

It will be appreciated if you put your Reference Number and your name as your paper revision file name (e.g. 001_TICATE 2020_Wulan).

You need to send us your revised manuscript in Microsoft Office Document file format (doc or docx) to the TICATE 2020 committee (ticate@untar.ac.id) by 31 July 2020 to avoid unnecessary delay.

Please complete your registration before 31 July 2020 for regular registration deadline. You are eligible to complete the payment before submitting the revision.

We invite you to present your paper at the conference. All the paper presented in TICATE 2020 will be published in IOP Conference Series: Materials Science and Engineering, and will be given to Scopus for indexing. All the publication decision will be on publisher site.

Further updated information will be published in our website (http://ticate.untar.ac.id)

If you have any questions, please do not hesitate to contact us. Thank you very much for your

cooperation.

Sincerely,

Dr. Hugeng, S.T.,M.T.(SMIEEE)
        Chairman