Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 8, July 2021: 2817-2823

Application of NLP:Design of Chatbot for New Research Scholars

Dr. Sunanda Mulik[a], Dr. Poonam Sawant[b], Dr. Vaishali Bhosale[c]

[a] Assistant Professor, Department of Computer Science, Vishwakarma University Pune, India
[b] Assistant Professor, Department of Computer Applications, SIMCA, Pune, India
[c] Assistant Professor, Department of Computer Science, Shivaji University Kolhapur, India

Abstract

Development of Intelligent Conversational Agent using Artificial Intelligence and Machine Learning technique is an interesting problem in the field of natural language processing (NLP). Conversational agents, widely known as chatbots, are being used as virtual assistant in many different domains including business, healthcare and government organizations in attempt to improve the quality of service provided.

This paper aims at using NLP towards the development of chatbot to assist the new research scholars. New research scholars often find it difficult where to start, how to start and have some common queries about fundamental concepts in research, research publications, data sources, funding agencies, indexing services etc. The paper addresses the need of virtual assistant to the new researcher scholars and describes the design of the chatbot model that would help them get the answers to the research queries at their fingertips. The database used in the study is generated by collecting data from web as well as through personal interactions with new research scholars. We have designed retrieval based text chatbot using Keras sequential model with Adagrad optimizer that is optimized for learning rate. The research also reveals the relationship between learning rate and accuracy for five optimizers. It was found that accuracy of Adam and RMSprop optimizers is inversely proportional to the learning rate whereas the accuracy of Adadelta optimizer is directly proportional to the learning rate. The model designed in this work is simple, computationally inexpensive and has succeeded to provide answers to user queries satisfactorily. The paper gives important directions in which research can further be extended.

Keywords: NLP, Chatbot Assistant, Retrieval Based, Sequential Model, Context Maintenance, Spell Check.

1. Introduction

Artificial Intelligence (AI) techniques are playing an important role in everyday life now. AI has made life very easy and comfortable (M. Dahiya (2017). These techniques include techniques like machine learning, deep learning, natural language processing etc. AI applications encompass a variety of applications from many different domains. These applications range from prediction systems to sentiment analysis. Dialogue Generation or Intelligent Conversational Agent development using Artificial Intelligence or Machine Learning technique is an interesting problem in the field of natural language processing (S. Ghose, J. J. Barua 2013). Conversational agents, also known as Chatbots, make conversation possible 24/7 and can be available in both the forms, text as well as speech. Chatbots can provide useful services, refer important information and present websites of interest to

users in addition to directly answering the question. hus, the users can get the intended information with little effort, or the conversation can be guided by the chatbot in a certain direction.

Chatbots are used as virtual assistants that do some routine tasks such as making calls, sending messages, booking appointments etc. for the user, or can be used as question- answering machines or for chatting purposes (R. P. Schumaker, H. Chen 2007). Conversation agents are predominantly used by businesses, government organizations and non-profit organizations. They are frequently deployed by financial organizations like bank, credit card companies, businesses like online retail stores and start-ups (Anjana Tiha 2018). There are also several non-commercial chatbots which are developed primarily for the purpose of fun and entertainment (Claus Mobus 2006). Other applications of chatbots include library services (Allison, DeeAnn 2011). flood assistance (Vinothini Kasinatha 2020). With the ever increasing awareness of technology, chatbot is becoming an everyday stuff now.

Based on the capability of input handling, chatbots are categorized as text based or speech based. Text based chatbots handle only text based communication whereas speech based chatbots are able to handle oral communication and are available in the speech form. Chatbots can also be categorized as Retrieval based or Generative, based on the way they produce the response. The goal of retrieval-based chatbot is to build a model that can retrieve the most appropriate response to a conversational input from a pool of candidate responses. Retrieval based bot uses a repository or database of predefined responses that is used to answer the user queries. In this type, the bot is trained to rank the best response from the finite set of responses (Deshpande Aditya 2017). This type of bot seldom makes mistakes and has higher accuracy as it is completely based on predefined responses. However, the same feature leads to some limitations also. The responses are too rigid and may not seem human or natural. As opposed, the Generative chatbot does not use any kind of predefined repository. This kind of bot generates a response by itself by using natural language generation techniques. It is trained using a database of past conversations, based upon which responses to the user are generated. Generative chatbots can be created using RNN and attention mechanisms techniques (Ramakrishna Kumar, Maha Mahmoud Ali (2020).

In this article, we explore the need of conversational agent to meet the information needs of new research scholars. New research scholars always find it difficult where to start, how to start and have some common queries about research , research publications, data sources, funding agencies, indexing services, impact factors and citation indices etc. They often spend much of their time in getting answers to these questions surfing on web or searching in books. In this paper we have developed a chatbot that would assist research scholars to get their basic research queries answered promptly and satisfactorily. Thereby, saving their time which can be utilized to carry out actual research work resulting in better productivity.

2. Theoretical Background

Neural networks, also known as artificial neural networks (ANNs) are a subset of machine learning and are at the heart of deep learning algorithms. ANN consists of node layers, containing an input layer, one or more hidden layers, and an output layer. Each layer consists of a number of artificial neurons called as nodes that are connected to the nodes in next layer. The connections have associated weight.

Net input to the neuron is computed using the following formula.

$$\sum_{i=1}^{m} w_i x_i + bias = w_1 x_1 + w_2 x_2 + w_3 x_3 + bias$$

Where, wi is the weight of the connection and xi is the input.

Each node consists of a threshold value and an activation function that produces node output. If the output of the node is above the specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data is passed along to the next layer of the network.

In this work, we have designed a simple neural network for the development of chatbot.

3. Material and Methodology

Database

The research uses question answer database that is generated by collecting data from web as well as through personal interactions with new research scholars. The data is stored in JSON format in which conversational intents are defined. Each conversational intent contains tag, pattern, responses and context information. The dataset contains data about more than two hundred research questions.

Platform Used

Operating System: Windows 10

Python Package: Anaconda3.5

Model Building Process

1) Building Vocabulary

Intent patterns are processed to build a vocabulary to be used in a model. Python WordNetLemmatizer is used to collapse distinct word forms which are then translated into bag of words.

2) Model Specification

Keras deep learning library is used to build a classification model. Keras sequential model is built that consists of three layers, input, hidden and output layer. Popular ReLu activation is used for input and hidden layer whereas output layer uses softmax activation function. The problem is about intent classification where classification output is a multiclass array, which would help identify encoded intent. Adagrad optimizer with learning rate of 0.05 is used after experimenting with other optimizers. Categorical_crossentropy function is used to measure the loss. The model is evaluated for Accuracy metric.

3) Experiment with different Optimizers at different learning rates

Optimizer and learning rate are the hyperparameters of the sequential model. The choice of optimizer and learning rate significantly affect the performance of the model. In the present work we attempted five optimizers namely Stochastic Gradient Descent (SGD), Adagrad, Adadelta, Adam and RMSprop

at different learning rates in order to optimize the accuracy of the model. The model is implemented with these five optimizers and the accuracy of the model is measured at five learning rate values of 0.001, 0.005, 0.01, 0.05 and 1. The results are shown in Table 1.

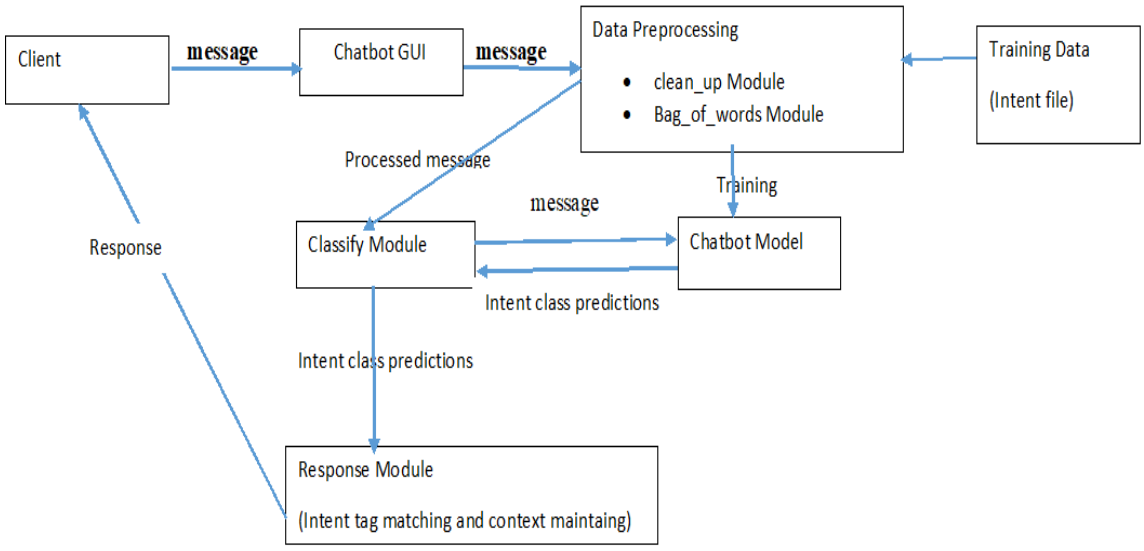
Table.1. Accuracy of Model for different Optimizers at different learning rates

Optimizer → Learning rate ↓	**SGD**	**Adagrad**	**Adadelta**	**Adam**	**RMSprop**
0.001	0.633	0.1651	0.0183	0.9633	0.9358
0.005	0.9174	0.6697	0.0734	0.9266	0.945
0.01	0.9358	0.8165	0.1376	0.8899	0.9358
0.05	0.1101	0.9908	0.4862	0.1193	0.3486
1	0.0459	0.1651	0.9633	0	0.0275

Chatbot Framework: Module Description

We have developed Retrieval Based text chatbot. The chatbot has a repository of predefined responses that is used to answer the user queries. The bot does the tasks of intent classification, entity identification and response selection. Based on the user query, an appropriate response will be selected from the repository and delivered to the user. Figure 1 shows the framework of the chatbot.

Figure.1. Chatbot Framework



1) Chatbot GUI Module

In this module, simple GUI with which user interacts with the bot is created. User can put his query in the query box. Spelling mistakes in user input are often a big hurdle in the success of any chatbot that leads to under usage of the bot (Bertsimas Jeton Arifi, Markus Ebner, Martin Ebner 2019). The present model provides spell check feature to autocorrect spelling mistakes in user input. Python enchant library is used to correct spelling errors in user input to enhance the responsive power of the bot. Before sending user input to the clean_up module, the input is checked for possible spelling mistakes and

corrected if any. The user query and the bot response is displayed in the chat window. After adding spell check feature, the performance of the bot is significantly increased.

2) clean_up module

Data cleaning is the first step in data pre-processing. The user input is often noisy and needs to be cleaned up. The clean_up module cleans up the input sentence and identifies the tokens. The identified tokens are then lemmatized to find the root words taking context into consideration.

3) bag_of_words Module

Algorithms cannot process data in a text form directly. The text needs to be broken down into a numerical form that the machine can process. This module translates sentence words into a vector representation called as a bag of words. It is a 0/1 array where each word is represented by digit 0 or 1 based on its presence or absence in the chatbot vocabulary. The bag of words is then fed to the classify module for further processing.

4) Classify Module

Classify module takes bag of words and finds its probability of belongingness to each intent. It returns the list of intents satisfying a predefined error threshold to the response module.

5) Response Module

Response module takes the intents from classify module. It performs two important tasks.

1. Context Maintaining

In any kind of conversation, context is very important to keep the flow of conversation and maintain the context from one turn of the conversation to the next, by maintaining the state of a conversation. Maintaining context is very easy in case of human conversation. But, it is often found difficult in chatbots to do so. Maintaining a context of conversation is one of the hardest problems faced by chatbot developers and most of the chatbot development frameworks provide an opinionated take about it.

The response module sets and checks for any context information available for the user. We have used dictionary data structure to store the state of the context. State is maintained in the form of unique identifier assigned to each user that helps the state-machine to maintain state for multiple users simultaneously. Context flow is defined in the intents. As soon as the intent is classified and backend logic finds a start of the context, the bot enters into the loop and ask related questions.

2. Tag Matching

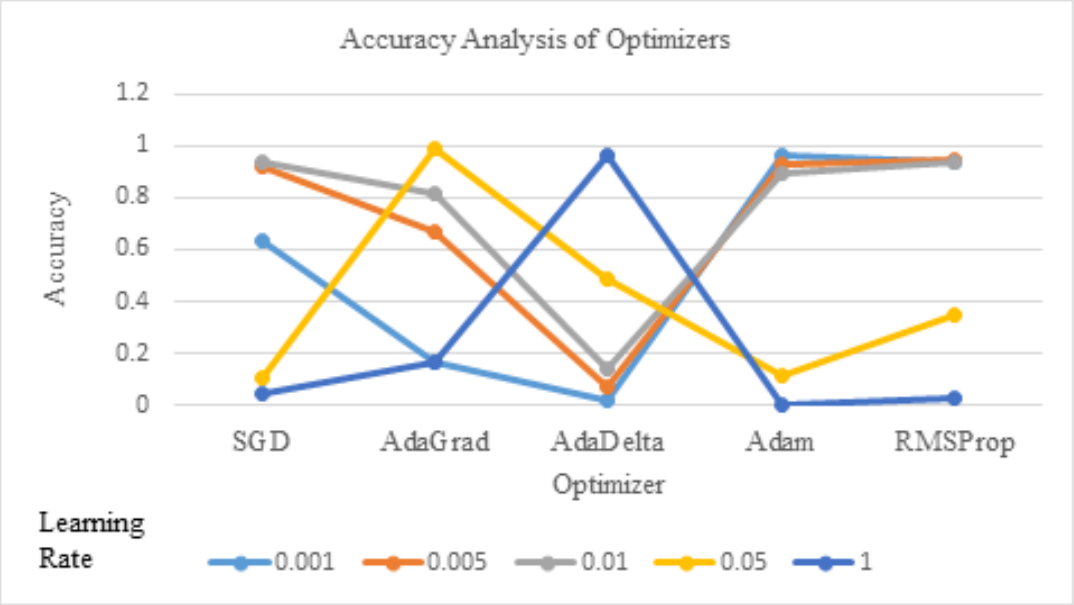
Response module checks for the matching tag in intents database. Once the matching tag is found, it selects response randomly from the list of available responses and returns it to the bot. If matching tag is not found, an empty response is returned.

4. Findings

The Figure 2 shows the graph of accuracy analysis for the five optimizers at five learning rates. The following observations are made from the graph.

- Accuracy of Adam and RMSprop optimizers is inversely proportional to the learning rate.
- Accuracy of Adadelta is directly proportional to the learning rate.
- SGD and Adagrad optimizers have higher accuracy at medium values of learning rate.

Figure.2. Graph of Accuracy Analysis of Optimizers



From the graph it is seen that, Adagrad optimizer has the highest accuracy amongst all at the learning rate 0.05. Hence, the Adagrad optimizer is used in final model building.

5. Discussion and Conclusion

Keras Sequential model is implemented for five different optimizers at different learning rates in order to analyze their impact on the performance of the model. The results obtained showed that Adagrad optimizer has better performance in terms of accuracy as compared to other four. The experimental results also revealed the relationship between learning rate and accuracy of the optimizers. It was found that accuracy of Adam and RMSprop optimizers is inversely proportional to the learning rate whereas the accuracy of Adadelta optimizer is directly proportional to the learning rate.

The Spell check feature used has increased the performance of the bot significantly. The model has succeeded in maintaining context and thus continue the flow of conversation to considerable extent that results in the increased satisfaction of the users.

The model design is very simple and computationally inexpensive. It is capable of answering basic queries of the new researchers satisfactorily in no time saving their time and efforts significantly. Features like spell check and context maintenance make it more responsive and increase its utility.

6. Limitations and Future Work

The current model suffers from few limitations that can be overcome. Following are some areas in which research can further be extended.

The present model uses auto-correct spell check feature that get the user input and then corrects any possible mistakes in the input. We can also add feature to provide spell suggestions when the user is typing a message.

- The current chatbot is retrieval based that is rigid and constrained to predefined, fixed set of responses. We think of extending the work towards development of generative chatbot that can generate the response based on user input.
- People usually prefer audio communication instead of written communication. The present chatbot is text based. It can be turned into speech based to increase its utility.
- People often find using mobile apps or messengers more convenient than web app. The future work aims at deploying the model as mobile application or whatsapp messenger to increase usage of the bot.

References

[1]. M. Dahiya (2017). "A Tool of Conversation: Chatbot," International Journal of Computer Sciences and Engineering, (IJCSE), 5(5), 2347-2693

[2]. S. Ghose, J. J. Barua (2013). "Toward the Implementation of a Topic Specific Dialogue Based Natural Language Chatbot as an Undergraduate Advisor," *International Conference on Informatics, Electronics & Vision*, 1-5

[3]. R. P. Schumaker, H. Chen (2007). *"*Leveraging Question Answer Technology to Address Terrorism Inquiry," *Decision Support Systems, 4(3), 1419-1430.*

[4]. Anjana Tiha (2018). "Intelligent Chatbot using Deep Learning", M.S. thesis, Dept. Comp. Sci.,Memphis Univ., Memphis.

[5]. Claus Mobus(2006). "Web communication with OpenSource chatbots, virtual fairs, rich media content*," Springer-Verlag Berlin Heidelberg.* https://doi.org/10.1007/3-540-29093-1.

[6]. Allison, DeeAnn (2011). "Chatbots in the Library: is it time?" Faculty Publications, UNL Libraries. 280. Available: https://digitalcommons.unl.edu/libraryscience/280.

[7]. Vinothini Kasinatha(2020). "FLOW: INTELLIGENT FLOOD AID CHATBOT AS VIRTUAL FLOOD ASSISTANT," *International Journal of Management*, 11(10) 1749-1756.

[8]. Deshpande Aditya (2017). "A Survey of Various Chatbot Implementation Techniques," *International Journal of Computer Engineering and Applications*, 11.

[9]. Ramakrishna Kumar, Maha Mahmoud Ali (2020). "A Review on Chatbot Design and Implementation Techniques," *International Research Journal of Engineering and Technology (IRJET)*, 7(2), 2791-2800.

[10]. Bertsimas Jeton Arifi, Markus Ebner, Martin Ebner (2019). "Potentials of Chatbots for Spell Check among Youngsters," *iJAI*, 1(1), 77-88.

[11]. B. P. Kiptonui (2013). "Chatbot Technology: A Possible Means of Unlocking Student Potential to Learn How to Learn," *Educational Research. Journal of Computer and System Sciences*, 74(3), 394–403.

[12]. J. Jia (2003). "The Study of the Application of a Keywords-based Chatbot System on the Teaching of Foreign Languages", Report of University of Augsburg, Augsburg, 1-36.

[13]. Garvit Bajpai, Rakesh Kumar Kannaujiya (2018). "CHATBOT IN PYTHON", B Tech. thesis, Dept. IT, APJ AKT Univ., Luckhnow, UP.

[14]. Ward, D (2005). "Why Users Choose Chat: A Survey of Behavior and Motivations", Internet Reference Services Quarterly, 10(1), 29-46.