

Chatbot using TensorFlow for small Businesses

¹Rupesh Singh
(Student)

Department of Computer Engineering
K J Somaiya College of Engineering
Mumbai, India
rupesh.singh@somaiya.edu

³Manmath Paste
(Student)

Department of Computer Engineering
K J Somaiya College of Engineering
Mumbai, India
manmath.p@somaiya.edu

⁵Nirmala Shinde
(Assistant Professor)

Department of Computer Engineering
K J Somaiya College of Engineering
Mumbai, India
nirmala.shinde@somaiya.edu

²Harshkumar Patel
(Student)

Department of Computer Engineering
K J Somaiya College of Engineering
Mumbai, India
harshkumar.jp@somaiya.edu

⁴Nitin Mishra
(Student)

Department of Computer Engineering
K J Somaiya College of Engineering
Mumbai, India
nitin.km@somaiya.edu

Abstract— Chatbots are software used in entertainment industry, businesses and user support. Chatbots are modeled on various techniques such as knowledge base, machine learning based. Machine learning based chatbots yields more practical results. Chatbot which gives responses based on the context of conversation tends to be more user friendly. The chatbot we are proposing demonstrates a method of developing chatbot which can follow the context of the conversation. This method uses TensorFlow for developing the neural network model of the chatbot and uses the nlp techniques to maintain the context of the conversation. This chatbots can be used in small industries or business for automating customer care as user queries will be handled by chatbots thus reducing need of human labour and expenditure.

Keywords— Chatbots; Neural Networks; Machine learning; Conversational models; Machine translation; Artificial intelligence; Statistics; Computational linguistics; Natural language processing; Text mining; TensorFlow; Tflearn.

I. INTRODUCTION

Chatbots are simple computer programs that interacts with users using natural language. The interaction can be textual or auditory depending upon the need. The popularity of this bots are increasing every day and are used in various practical applications which include customer service, information acquisition and dialogue systems. This bot uses Natural language processing techniques for understanding user query and give response. Now there are bots which uses fixed response techniques where user query is searched for fixed keywords and returns response which matches the most with the query. Also, some bots use string manipulation techniques

where response is nothing but manipulated version of user query e.g. Eliza.

Small businesses and small startups generally have a high customer to employee's rate. They cannot attend to every customer personally. Such small businesses need automation in their customer support system. Chatbot can automate the customer support part. Chatbot can increase the profits by decreasing the resources strain. We here provide the method for making such chatbots that a small company can make with very little resources. This can be useful for many other organizations such as universities and colleges, as displayed in our demo.

In this project, we created a chatbot which will provide information about admission in K.J. Somaiya college of engineering to the student through web interface using TensorFlow and NLP techniques. Thus, our bot will act as a person at enquiry office whom they can ask any questions regarding admission and college environment at any time in more interactive way. Thus, reducing workload of the admission cell personnel. This system can be integrated on college website and thus students who desires to get admission in the college can clear their doubts from their home. we have provided a feedback mechanism through will user can rate bot's responses and thus bot can respond more accurately next time.

TensorFlow is used to create Neural model on which bot will be trained based on intent file. This model can be updated as and when new rule is implemented in the college. This bot can maintain context and can give responses based on context. Thus, this paper focuses on the sample bot which can be used as a reference for creating chatbots for many other purposes. The paper talks about the methods and technologies used in

creating this chatbot, and why some things are preferred over other. The paper by explaining the inner working of this sample chatbot explains the proposed methodology of developing similar purpose chatbot.

In the following sections we discuss:

1. Existing chatbots and there strong and weak points.
2. How the proposed chatbot can act as a reference for creating similar bots for small businesses.
3. Inner working of the proposed method and algorithms used.
4. Result of the proposed system and comparison with other methods.

II. RELATED WORK

Few Chatbots with useful applications are presented. We will discuss about ELIZA and ALICE as of now.

ELIZA

The first chatbot developed was ELIZA [1]. It was developed by Joseph Weizenbaum using a keyword matching and string manipulation techniques. When user provides the query, this chatbot searches for certain keyword. These keywords were already defined and if found then the response was retrieved. If not found, then Eliza would perform certain string manipulation techniques or use certain rules to obtain detail knowledge from the user to continue the interaction.

Consider the following example to understand how this would work.

USER: How can I get in touch with my mother?

Clearly ‘Mother’ is the keyword in the given user query, so the bot will try to find words related to mother like family. Then by using some string operation it will give an answer.

ELIZA: Tell me more about your family

Also, Eliza uses string manipulation techniques where it replaces some words from user query and gives it as a response.

USER: I am fine

ELIZA: How long have you been fine?

The chatbot doesn't understand what it is saying. It only produces results according to rules, static responses and string manipulation. However most of the time there are some

standard replies that ELIZA would use if a keyword is not found. Some examples would be “Very interesting, please go on” or “Can you think of a special example?”.

ELIZA: How do you do. Please tell me your problem.
YOU: i need your help
ELIZA: Do you want my help ?
YOU: i have to complete my paper
ELIZA: Your paper ?
YOU: yes, research paper
ELIZA: I understand.
YOU: i am sad and very tired to complete it
ELIZA: Do you think coming here will help you not to be sad ?
YOU: people told me you are a doctor so
ELIZA: Do you sometimes wish you were a doctor so ?
YOU: i want to complete my project as well
ELIZA: Why do you say your project as well ?

Talk Reset

Figure 1: Chatting session with Eliza.

ALICE

ALICE was implemented by Richard Wallace in 1995 [6]. Eliza was source of inspiration for Alice. It is a natural language processing bot which uses pattern matching for retrieving responses. The information is stored in Artificial Intelligence Markup Language also known as AIML files. AIML files were specifically designed to store pattern knowledge base for chatbots. These files are like XML files. AIML are of three categories.

Atomic categories: [6]

In this AIML category there is exact match and thus response is generated directly.

```
<category>
<pattern>What is your name? </pattern>
<template>My name is john</template>
</category>
```

Default categories: [11]

Default categories are general form of input query which makes use of wild character such as * so as to match user submitted query.

```
<category>
<pattern>What is *</pattern>
<template>It is my name</template>
</category>
```

If response is not generated using first category it will proceed to this category. Consider if user enters, “what is john?”, Alice will search for default category and if match is found response is displayed.

Recursive categories: In this category special tags are used to specify recursion and get more information from user [8,9].

```
Rupesh: hello
ALICE: Hi there!
Rupesh: i need help
ALICE: Can you ask for help in the form of a question?
Rupesh: can you have me?
ALICE: No I don't think I have any you. But I do have a collection of
gossip.
Rupesh: tell me gossip
ALICE: Hal said he translate binary code into a language and
communicate the language to me.
Rupesh: my name is rupesh
ALICE: I am always glad to make new friends, Rupesh.
Rupesh: do you know my name?
ALICE: You said your name is Rupesh?
Rupesh: how long will you remember my name
ALICE: I will store your name in a hash table until the next time my
program restarts. After that, it will remain in my long-term log files.
```

Figure 2: Chatting session with Alice

III. PROPOSED SYSTEM

The proposed system uses the machine learning at its core. The proposed system uses the TensorFlow to make a neural network and train it with intent file to generate a response model. This response model can be used to predict the response from the query of the user. The proposed system consists of three main parts

- 1) User Interface
- 2) Neural network model and NLP unit
- 3) Feedback System

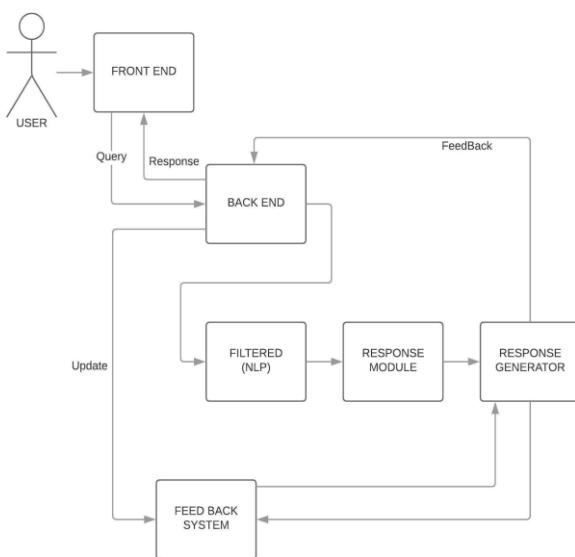


Figure 3: Architecture

1. User interface

The user interface can be in the form of anything from desktop to phone and web. The interface in the example is a web interface. The interface will also work as the instruction guide for the users. The user will get all the information needed to understand the system in here. The main purpose of the interface will be to take the queries and convey them to the backend which consist of tensor module and feedback system. The interface works as a medium between the backend and user.

2. Neural Model and NLP Unit

This module is the main core of the whole system. This is the part which generates the actual response for the user query. First the model is defined using TensorFlow which is trained using intent file created. Intent file is in form of json file which is as follows:

```
{"tag": "admission_enquiry",
"patterns": ["i have some doubts regarding process", "I want
to know about admission process", "enquiry about
admission"],
"responses": ["Okay, what do you wanna know?", "what are
you doubts?", "Okay, you can ask"]
},
{"tag": "application_date",
"patterns": ["when do i apply?", "What is the due date for
application?", "what is last date of the application?", "what
is starting date of the application?" ],
"responses": ["Application process will remain open at 04th
of July to 14th of august", "from 4th July to 14th
august", "you can apply from 4th july to 14th august"]
},
{"tag": "Application_form",
"patterns": ["how to avail admission form?", "From where to
have admission form?", "where can I get an application
form? "],
"responses": ["Application form for admission is available at
the Somaiya website", "from somaiya website", "you can get
forms from somaiya website", "visit somaiya website for
forms"]
},
{"tag": "Location",
"patterns": ["where is the location of the college?", "College is
located at?", "How to reach college?" ],
"responses": ["Somaiya college of engineering is Located
near vidyavihar station at central railway.", "College is
located in mumbai near vidyavihar railway station."]
},
```

It consist of three parts i.e. Tags, Patterns , Responses. The tag is nothing but context of that query. It defines what query is about. Patterns and responses as name suggests is used to train the model with sentences and get corresponding responses. Patterns are loaded and passed through query modulation

process. Here NLP takes place where various functions are applied in form of pipeline which includes [2] Sentences -> Tokenization -> Lemmatization -> POS-tagging. This data is then stored in form of bag of words which is used as input to training model. We have used bag of words technique for feature extraction. Feed forward network is created using tensorflow with 4 layers (1 Input layer + 2 hidden layer + 1 output layer). Trained model is saved which is used for predicting responses taking input as user query.

During training, the model creates a bag of words array which is collection of all unique words. Now when query is passed, it goes through query modulation process and converted to bag of words. Here the words which are present is represented by 1 while others as 0.

Consider the query

QUERY: how to get admission in college?

Following is bag of words representation for the given query:

Now when user gives query this query will be passed through same query modulation process and converted into bag of words. This will be given as input to model and model will return all the tags that have probabilities greater than error threshold defined. Then response model will select the response which will be more accurate and having greater ranking given by feedback System.

QUERY: Hello
PREDICTION: [('greeting', 0.98416048)]
RESPONSE: Hello, thanks for visiting

Here model predicted intent greeting with probability of 0.98

QUERY: How to get admission in college?
PREDICTION: [('Admission_process', 0.99881846)]
RESPONSE: For admission person should follow the given steps....

Similarly here model predicted intent Admission process with probability of 0.9988

QUERY: Any hostel facilities available
PREDICTION:[('Hostel_availability', 0.63436478), ('greeting', 0.16629744)]
RESPONSE: There are two hostels available for students

In this case model predicted two intent Hosel_availability with probability of 0.6343 and greeting with probability of 0.166 as Threshold is 0.15 as of now. The intent with higher probability is selected and response is generated accordingly.

QUERY: activities related to sports

PREDICTION: [('Sport_support', 0.99963164)]

RESPONSE: Skream is a national.....

3. Feedback System

The database stores the intent file which is crucial part of the training and response. User interface will be provided with functionality through which user can rate the response and this will be stored in database along with responses. Next time when user gives query, response model will give response based on the feedback i.e. this response will be checked for ranking and highest-ranking response will be displayed. This ranking will be integer based with 5 having highest ranking. Initially all the responses will be set to default 0 and as and when user interacts with the bot it will be updated.

IV. RESULTS

We have used machine learning approach to create a bot in this paper. Chatbots based on machine learning does not understand the meaning of sentences. It learns how to respond based on previous experience. Though we have used some NLP functions but the actual process through which response is generated is using machine learning. As said earlier, we created the model and trained it with the intent file thus more diverse the intent file more accurate will be the result. This model is used to predict the tag based on the user query.

Now coming to the responsiveness of the proposed chatbot which is response generation and maintaining the context of the conversation, the bot performs quite efficiently if the query asked is simple and in accordance to the intent file created. Also, context is maintained to some extent. For checking the efficiency of the system, we conducted a small experiment where we asked 10 people to interact with the bot and give their feedback. These people were students and their parents who went through the admission process and have to wait in long queue in the college just to get their doubts cleared. The results are plotted in the following graph.

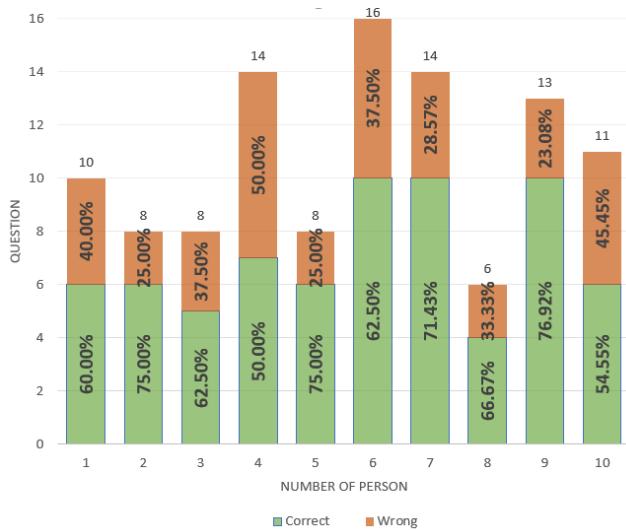


Figure 4: Output Graph

Each person asked different number of question depending upon their requirement. The graph plots the satisfactory (Correct) vs unsatisfactory(wrong) responses. for eg, person1 feels 60% of the responses were correct while 40% were wrong. Thus taking the whole experiment into consideration, total 108 questions where asked out of which 74 where satisfactory which gives accuracy of 68.51%. This accuracy can further be enhanced with more varied intent file.

	form		0.99642229]
5	who is somaiya college principal	Principal of somaiya college of engineering is XXXX XXX	[('Who_Principal', 0.4513315), ('Founder', 0.40391245)]
6 *	how to get admission form	For admission person should follow the given steps.....	[('Admission_process', 0.7114048), ('Application_form', 0.16649978)]
7	where is college located	Somaiya college of engineering is Located near vidyavihar station at central railway.	[('Location', 0.86157036)]
8	what payment modes?	we do not accept cash, but net banking , card payment will do	[('payments_modes', 0.98774439)]
9	mode of payment	Net banking.	[('payments_modes', 0.80774439)]

TABLE 1: RESULTS

S R. N O	QUERY	RESPONSE	PROBABILI-TY
1	hii	Hello, thanks for visiting	[('greeting', 0.24204822), ('Location', 0.2273401)]
2	enquiry about admission	what are you doubts?	[('admission_enquiry', 0.99886227)]
3	what eligibility criteria for admission	basic require- ment is 12th (provided you clear the cut off)	[('Eligibility_nationality', 0.99982148)]
4	From where to get application	visit somaiya website for forms	[('Application_form',

Following are some of the interactions performed and prediction given. On an average, out of 10 queries asked the bot give 6-7 responses correct. But again, it depends upon the way questions are framed. It is assumed the user has good English knowledge and can interact with the system properly. With increase in the dataset this accuracy can further be enhanced, and system can be made more efficient

V. CONCLUSION

Main objective of this project was to provide a system to create chatbot which can be used by small businesses as a replacement of customer support.

The method demonstrated was successful to some extent in creating chatbots when the domain is small but, the result got will get more accurate as the size of intent file increases. The accuracy of the chatbot is directly proportional to the size of intent file used for training the chatbot. With small domain it is relatively easy to create the intent files that will yield a certain level of accuracy. This method is clearly suitable in situation where the domain is narrow, and user interact with some relevance.

VI. FUTURE SCOPE

The method presented uses solely machine learning with some helping natural language processing for converting the intent into bag of words. The method can be improved by using some additional level natural language processing. The natural language processing can be used to do sentiment analysis which can complement the machine learning predictor. The sentimental part can be analyzed using natural language processing which will emphasize on the keywords in the user query which can be very useful in reducing the error in prediction.

Instead of using natural language processing to perform the sentimental analysis and find out the key aspect of queries we can add multiple machine learning neural models to do the same. The first method is quite hard to achieve as it is very hard to determine the keywords and emphasis in query using natural language processing. The machine learning does not understand sentimental analysis as it is purely based on mathematical functions.

- [8] B.A shawar and E. Atwell, "Evaluation of Chatbot Information System, in Eighth Maghrebian Conference on Software Engineering and Artificial Intelligence.
- [9] B.A shawar and E. Atwell, "JLCL,"2007. [Online]. Available: http://www.jlcl.org/2007_heft1/Bayan_Abu-shawar_and_Eric_Atwell.pdf
- [10] B. Setiaji and F. W. Wibowo, "Chatbot Using a Knowledge in Database: Human-to-Machine Conversation Modeling," in Intelligent \ Systems, Modelling and Simulation (ISMS), 2016 7th International Conference on, Bangkok, Thailand.
- [11] B. R. Ranoliya, N. Raghuwanshi and S. Singh, "Chatbot for university related FAQs," in Advances in Computing, Communications and Informatics (ICACCI), 2017.

REFERENCES

- [1] A. Vichare, A. Gyani, Y. Shrikhande and N. Rathod, "A chatbot system demonstrating Intelligent Behaviour using NLP," in *International Journal of Advanced Research in Computer Engineering & Technology (IJARCE)*, 10 October 2015. J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp. 68-73.
- [2] R. Shah, S. Lahoti and P. L. K., "An Intelligent Chat-bot using Natural Language Processing," in *International Journal of Engineering Research*, 1 May 2017. T. L. Gilbert, *Formulation, Foundations and Applications of the Phenomenological Theory of Ferromagnetism*, Ph.D. dissertation, Illinois Inst. Tech., Chicago, IL, 1956, unpublished.
- [3] A. Sordoni, M. Galley, M. Auli, C. Brockett, Y. Ji, M. Mitchell, J.-Y. Nie, J. Gao and B. Dolan, "A Neural Network Approach to Context-Sensitive Generation of Conversational Responses," in *arXiv:1506.06714*, 22 June 2015.S. O. Demokritov and V. E. Demidov, "Micro-Brillouin light scattering spectroscopy of magnetic nanostructures," *IEEE Trans. Magn.*, to be published.
- [4] "Contextual chatbots using tensorflow," 2017. [Online]. Available: <https://chatbotsmagazine.com/contextual-chat-bots-with-tensorflow-4391749d0077>.Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *IEEE Transl. J. Magn. Jpn.*, vol. 2, pp. 740-741, August 1987 [*Dig. 9th Annual Conf. Magn. Jpn.*, p. 301, 1982].
- [5] "Make bots great again," 30 September 2016. [Online]. Available: <http://blog.datatomic.com/2016/09/make-bots-great-again.html>.
- [6] "Cornell University Library," 2014. [Online]. Available: <https://arxiv.org/ftp/arxiv/papers/1408/1408.6762.pdf>.
- [7] "Text Classification using Neural Networks," 2017. [Online]. Available: <https://machinelearnings.co/text-classification-using-neural-networks-f5cd7b8765c6>.