

Project Synopsis

ON

CHATBOT IN PYTHON

Submitted for partial fulfilment of award for

Bachelor of Technology

In

Information Technology (2021)

Submitted by

Shivam Pandey (17122123069)

Prashant Kumar Upadhyay(1812213801)

Under the guidance of

Mr. Ajay Kr. Srivastava



**Shri Ramswaroop College of Engineering and
Management**

Lucknow

Affiliated to

Dr. APJ Abdul Kalam Technical University

1.1 ABSTRACT

A chatbot is a computer program that can converse with humans using artificial intelligence in messaging platforms. The goal of the project is to add a chatbot feature and API for Yioop. discussion groups, blogs, wikis etc. Yioop provides all the basic features of web search portal. It has its own account management system with the ability to set up groups that have discussions boards. Groups are collections of users that have access to a group feed. The user who creates a group is set as the initial group owner. Posts are grouped by thread in a group containing the most recent activity at the top. The chatbot API for Yioop will allow developers to create new chatbots, powered by rules or artificial intelligence, that can interact like a human with users in a groups feed page. Example chatbots that can be developed with this API is weather chatbots or book flight chatbots. Over past few years, messaging applications have become more popular than Social networking sites. People are using messaging applications these days such as Facebook Messenger, Skype, Viber, Telegram, Slack etc. This is making other businesses available on messaging platforms leads to proactive interaction with users about their products. To interact on such messaging platforms with many users, the businesses can write a computer program that can converse like a human which is called a chatbot.

Chatbots come in two kinds:

- Limited set of rules
- Machine learning

Chatbot that uses limited set of rules

This kind of bots are very limited to set of texts or commands. They have ability to respond only to those texts or commands. If user asks something different or other than the set of texts or commands which are defined to the bot, it would not respond as desired since it does not understand or it has not trained what user asked. These bots are not very smart when compared to other kind of bots.

1.2 Chatbot and Machine learning

Machine learning chatbots work using artificial intelligence. Users need not to be more specific while talking with a bot because it can understand the natural language, not only commands. This kind of bots get continuously better or smarter as it learns from past conversations it had with people.

Here is a simple example which illustrates how they work. The following is a conversation between a human and a chatbot: Human: "I need a flight from San Jose to New York." Bot: "Sure! When would you like to travel?" Human: "From Dec 20, 2016 to Jan 28, 2017." Bot: "Great! Looking for flights."

In order to achieve the ultimate goal, I have taken an iterative approach and divided my work into four major deliverables. These deliverables not only helped me in understanding the code structure of Yioop but also enhance Yioop's functionality. In the rest of the report, I will be discussing about the four deliverables. To understand more on chatbot service, I had implemented a Facebook Messenger Weather Bot in deliverable 1, which is discussed in next section. The purpose of deliverable 2 is to introduce chatbots to the Yioop. I have added Bot Configuration settings which is used to add bot users in Yioop. In the next deliverable, I have added a functionality where the user will be able to call bots in a group thread. Activation of bots will happen by calling respective callback URL which is already configured that helps bots to have a conversation with users. More details on this is discussed in deliverable 3 section. As a deliverable 4, I have created a weather bot i.e., a web application in php that calls yahoo API to get weather information. The last section of the report contains the conclusion and future work.

I have implemented a Facebook Messenger Bot to get an overview of how chatbot is built. During this implementation, I understood the flow of control for a chatbot service with other services which is explained below.

In order to create a Facebook Messenger Bot, a developer needs to be authenticated and approved by Facebook to converse with the public and the web server for security reasons. For a Facebook Messenger Bot, I have created a simple web application using Node.js by installing the necessary dependencies using npm. I ran this locally. I also downloaded and installed ngrok and started it - `npm run ngrok`. This launched a Forwarding URL to the local running server, that means any requests to Forwarding URL will hit the locally running server. This url is used as a

Callback URL in Facebook App which will be explained further. To set up the Facebook App, I have created a Facebook Page and Facebook App using my Facebook account. While setting up a Webhook in the app settings, I have given the Forwarding URL as Callback URL and added code for verification. The access token in page settings is stored as environment variable as it will be used in integration. In order to make webhook to receive messages from this page, the app is subscribed to the page created. To set up the bot to handle the POST calls at webhook, I have created a webhook endpoint in the sample application.

1.3 Artificial Intelligence

AI was coined by John McCarthy, an American computer scientist, in 1956 at The Dartmouth Conference where the discipline was born. Today, it is an umbrella term that encompasses everything from robotic process automation to actual robotics. It has gained prominence recently due, in part, to big data, or the increase in speed, size and variety of data businesses are now collecting. AI can perform tasks such as identifying patterns in the data more efficiently than humans, enabling businesses to gain more insight out of their data.

Types of artificial intelligence

AI can be categorized in any number of ways, but here are two examples.

The first classifies AI systems as either weak AI or strong AI. Weak AI, also known as narrow AI, is an AI system that is designed and trained for a particular task. Virtual personal assistants, such as Apple's Siri, are a form of weak AI.

Strong AI, also known as artificial general intelligence, is an AI system with generalized human cognitive abilities so that when presented with an unfamiliar task, it has enough intelligence to find a solution. The Turing Test, developed by mathematician Alan Turing in 1950, is a method used to determine if a computer can actually think like a human, although the method is controversial.

The second example is from Arend Hintze, an assistant professor of integrative biology and computer science and engineering at Michigan State University. He categorizes AI into four types, from the kind of AI systems that exist today to sentient systems, which do not yet exist. His categories are as follows:

Type 1: Reactive machines.

An example is Deep Blue, the IBM chess program that beat Garry Kasparov in the 1990s. Deep Blue can identify pieces on the chess board and make predictions, but it has no memory and cannot use past experiences to inform future ones. It analyzes possible moves -- its own and its opponent -- and chooses the most strategic move. Deep Blue and Google's AlphaGO were designed for narrow purposes and cannot easily be applied to another situation.

Type 2: Limited memory.

These AI systems can use past experiences to inform future decisions. Some of the decision-making functions in autonomous vehicles have been designed this way. Observations used to inform actions happening in the not-so-distant future, such as a car that has changed lanes. These observations are not stored permanently.

Type 3: Theory of mind.

This is a psychology term. It refers to the understanding that others have their own beliefs, desires and intentions that impact the decisions they make. This kind of AI does not yet exist.

Type 4: Self-awareness. In this category,

AI systems have a sense of self, have consciousness. Machines with self-awareness understand their current state and can use the information to infer what others are feeling. This type of AI does not yet exist.

Examples of AI technology

Automation is the process of making a system or process function automatically. Robotic process automation, for example, can be programmed to perform high-volume, repeatable tasks normally performed by humans. RPA is different from IT automation in that it can adapt to changing circumstances.

Machine learning is the science of getting a computer to act without programming. Deep learning is a subset of machine learning that, in very simple terms, can be thought of as the automation of predictive analytics. There are three types of machine learning algorithms: supervised learning, in which data sets are labeled so that patterns can be detected and used to label new data sets; unsupervised learning, in which data sets aren't labeled and are sorted

according to similarities or differences; and reinforcement learning, in which data sets aren't labeled but, after performing an action or several actions, the AI system is given feedback.

Machine vision is the science of making computers see. Machine vision captures and analyzes visual information using a camera, analog-to-digital conversion and digital signal processing. It is often compared to human eyesight, but machine vision isn't bound by biology and can be programmed to see through walls, for example. It is used in a range of applications from signature identification to medical image analysis. Computer vision, which is focused on machine-based image processing, is often conflated with machine vision.

Natural language processing (NLP) is the processing of human -- and not computer -- language by a computer program. One of the older and best known examples of NLP is spam detection, which looks at the subject line and the text of an email and decides if it's junk. Current approaches to NLP are based on machine learning. NLP tasks include text translation, sentiment analysis and speech recognition.

Pattern recognition is a branch of machine learning that focuses on identifying patterns in data. The term, today, is dated.

Robotics is a field of engineering focused on the design and manufacturing of robots. Robots are often used to perform tasks that are difficult for humans to perform or perform consistently. They are used in assembly lines for car production or by NASA to move large objects in space. More recently, researchers are using machine learning to build robots that can interact in social settings.

1.4 AI applications

AI in healthcare. The biggest bets are on improving patient outcomes and reducing costs. Companies are applying machine learning to make better and faster diagnoses than humans. One of the best known healthcare technologies is IBM Watson. It understands natural language and is capable of responding to questions asked of it. The system mines patient data and other available data sources to form a hypothesis, which it then presents with a confidence scoring schema. Other AI applications include chatbots, a computer program used online to answer questions and assist customers, to help schedule follow-up appointments or aiding patients through the billing process, and virtual health assistants that provide basic medical feedback.

AI in business. Robotic process automation is being applied to highly repetitive tasks normally performed by humans. Machine learning algorithms are being integrated into analytics and CRM platforms to uncover information on how to better serve customers. Chatbots have been

incorporated into websites to provide immediate service to customers. Automation of job positions has also become a talking point among academics and IT consultancies such as Gartner and Forrester.

AI in education. AI can automate grading, giving educators more time. AI can assess students and adapt to their needs, helping them work at their own pace. AI tutors can provide additional support to students, ensuring they stay on track. AI could change where and how students learn, perhaps even replacing some teachers.

AI in finance. AI applied to personal finance applications, such as Mint or Turbo Tax, is upending financial institutions. Applications such as these could collect personal data and provide financial advice. Other programs, IBM Watson being one, have been applied to the process of buying a home. Today, software performs much of the trading on Wall Street.

AI in law. The discovery process, sifting through of documents, in law is often overwhelming for humans. Automating this process is a better use of time and a more efficient process. Startups are also building question-and-answer computer assistants that can sift programmed-to-answer questions by examining the taxonomy and ontology associated with a database.

AI in manufacturing. This is an area that has been at the forefront of incorporating robots into the workflow. Industrial robots used to perform single tasks and were separated from human workers, but as the technology advanced that changed.

Background And Literature Review

2.1 Literature Review

Purpose	Papers	Year of publications	Pros	Cons
Usefulness of chatbots	Bayan Abu Shawar and Eric Atwell, 2007 "Chatbots: Are they Really Useful?"	2007	<ul style="list-style-type: none"> Significance of Chatbot 	Machine like Response
Natural language processing	Journal for Computational Linguistics and Language	2007	<ul style="list-style-type: none"> Benefit of natural language processing in any AIML chatbot Detailed overview of Natural language processing mechanism 	<ul style="list-style-type: none"> Technical and obsolete problems in Nlp mechanism
Real time solution for Open learner model	Towards natural language negotiation of open learner models	2009	<ul style="list-style-type: none"> Use of technology to enhance response. 	<ul style="list-style-type: none"> Limited information Problem in integration
Intelligent tutorial system	Whitepaper for the Army's Science of Learning Workshop, Hampton	2006	<ul style="list-style-type: none"> Detailed description about Python and AIML 	<ul style="list-style-type: none"> Only for chatbot based on Aimpl
Aimpl and Brain loading	AI Chat Bot with AIML Submitted by NanoDano	2015	<ul style="list-style-type: none"> Effective way of installation. 	<ul style="list-style-type: none"> Slow brain loading

Tab 2.1 Literature Survey

2.2 Natural Language Processing

Natural Language Processing (NLP) is the study of letting computers understand human languages[3]. Without NLP, human language sentences are just a series of meaningless symbols to computers. Computers don't recognize the words and don't understand the grammars. NLP can be regarded as a "translator", who will translate human languages to computer understandable information.

Traditionally, users need to follow well-defined procedures accurately, in order to interact with computers. For example, in Linux systems, all commands must be precise. A single replace of

one character or even a space can have significant difference. However, the emergence of NLP is changing the way of interacting. Apple Siri and Microsoft Cortana have made it possible to give command in everyday languages and is changing the way of interacting.

2.3 Machine Learning.

Machine Learning (ML) is an area of computer science that "gives computers the ability to learn without being explicitly programmed". The parameter of the formulas is calculated from the data, rather than defined by the programmer. Two most common usage of ML is Classification and Regression. As shown in figure1[8], Classification means to categorize different types of data, while Regression means to find a way to describe the data. Basic ML program will have two stages, *fitting* and *predicting*. In the fitting stage, the program will be given a large set (at least thousands) of data. The program will try to adjust its parameter based on some statistical models, in order to make it "fit" the input data best. In the predicting stage, the program will give a prediction for a new input based on the parameters it just calculated out. For example, the famous Iris flower dataset [9] contains the measurement of several features of three different species of flowers, such as the length of sepals and petals. A well-defined ML program can learn the pattern behind this feature and give prediction accordingly.

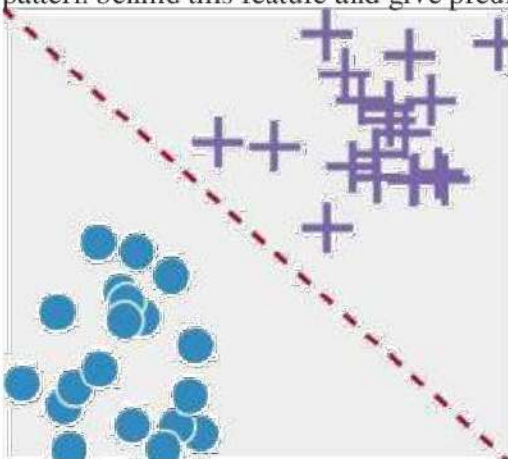


Figure2a: Classification

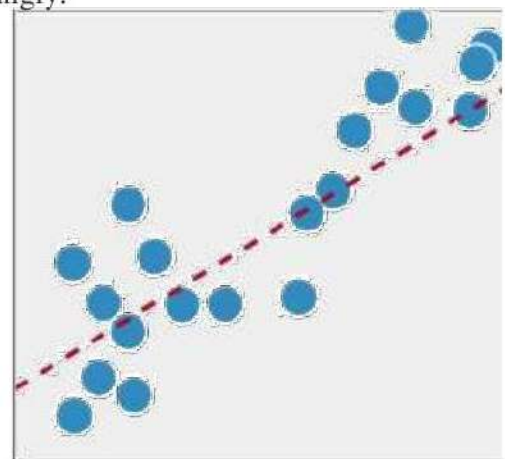


Figure2b: Regressing

PROBLEM IDENTIFICATION

3.1 Problem Identification-

Changing of requirements As an industrial project to build a product, we must follow the requirement from the user. However, because the project's goal is to be used by the business team, but it is responsible by the technical team, the requirement changed a lot in the middle after a meeting with the business team. The business team want a simple bot that can give recommendation immediately. We had to archive what we had done before and build a new one.

- **Lacking of training data.** The quantity and quality of training data is critical to the performance to a machine learning model. However, because some confidential and privacy reasons, the business team cannot provide enough data for us, and we had to make up data by our own. For the machine learning model, we generate some fake data based on our daily life experience, which is really biased, although with a good accuracy on the fake data.
- **Unstable API version.** Because API service we are using are still under development, and we cannot fix to a version for the API, the API may changes overtime. Moreover, there are inconsistencies between the APIs and their documents or sample codes.
- **Not familiar with the PHP language and .NET framework.** None of the three of us has previous knowledge with PHP language. Programming in a new language in such a huge framework is quite challenging for us at the beginning of the project. However, when we comes to the later phases, we are more used to that.

REQUIREMENTS AND SPECIFICATION

3.1 Hardware Requirements:

Processor	: Pentium IV(minimum)
Hard Disk	: 40GB
RAM	: 256MB (minimum)

3.2 Software Requirements:

Operating System	: Windows or Linux
Technology	: PYTHON, AIML

CONCLUSION AND FUTURE SCOPE

9.1 Conclusions

Chatbots are the new Apps! As we have discussed in the above deliverables, this project brings the power of chatbots to Yioop and enriches its usability. Chatbots in Yioop can give a human like touch to some aspects and make it an enjoying conversation. And they are focused entirely on providing information and completing tasks for the humans they interact with. The above mentioned functionality in all the deliverables is implemented and pushed in to Yioop code. By implementing the above mentioned deliverables I was able to add a basic chatbot functionality in to the Yioop. I.e., configuring and creating accounts for bot users with bot settings which is mentioned in deliverable 2, activating a bot whenever a user asks for it via post in a thread which is discussed in deliverable 3 and as I discussed in deliverable 4, I have implemented a simple weather chatbot that gives weather information whenever a user ask and Fig. 3 tells that I was also able to converse with the bot in Yioop. I intend to enhance the system developed so far in CS298. Next step towards building chatbots involve helping people to facilitate their work and interact with computers using natural language or using set of rules. Future Yioop chatbots, backed by machine-learning technology, will be able to remember past conversations and learn from them to answer new ones. The challenge would be conversing with multiple bot users and multiple user

9.2 Future Scope

There are limitations to what has been currently achieved with chatbots. The limitations of data processing and retrieval are hindering chatbots to reach their full potential. It is not that we lack the computational processing power to do so. However, there is a limitation on “How” we do it. One of the biggest examples is the retail customer market. Retail customers are primarily interested in interacting with humans because of nature of their needs. They don’t want bots to process their needs and respond accor

References

1. Bayan Abu Shawar and Eric Atwell, 2007 “Chatbots: Are they Really Useful?”
2. LDV Forum - GLDV Journal for Computational Linguistics and Language Technology.
3. http://www.ldv-forum.org/2007_Heft1/Bayan_Abu-Shawar_and_Eric_Atwell.pdf
4. Bringing chatbots into education: Towards natural language negotiation of open learner models. Know.-Based Syst. 20, 2 (Mar. 2007), 177-185.
5. Intelligent Tutoring Systems: Prospects for Guided Practice and Efficient Learning. Whitepaper for the Army's Science of Learning Workshop, Hampton, VA. Aug 1-3, 2006.
6. <http://en.wikipedia.org/wiki/Chatterbot>
7. ALICE. 2002. *A.L.I.C.E* AI Foundation, <http://www.alicebot.org/>