A
Project Report
On

AI POWERED CHATBOT

By

HARSHAL VINAYAK SHINDE

DEPARTMENT OF MCA (ENGINEERING)

Gokhale Education Society's
R.H.SAPAT COLLEGE OF ENGINEERING, MANAGEMENT STUDIES AND
RESEARCH, Nashik-5
[2020-2021]

A
Project Report
On

AI POWERED CHATBOT

In partial fulfillment of requirements for the degree of

Master of Computer Application
In
Engineering

Submitted By:

HARSHAL VINAYAK SHINDE

Under the Guidance of

Prof. TRUPTI U. AHIRRAO

DEPARTMENT OF MCA (ENGINEERING)

Gokhale Education Society's
R.H.SAPAT COLLEGE OF ENGINEERING, MANAGEMENT STUDIES,
AND RESEARCH, Nashik-5 [2020-2021]

Gokhale Education Society's
R.H.SAPAT COLLEGE OF ENGINEERING, MANAGEMENT STUDIES AND
RESEARCH, Nashik-5
[2020-2021]

<u>CERTIFICATE</u>

This is to certify that-

HARSHAL VINAYAK SHINDE
(Roll.No:22)

has completed his Project work and submitted the report on, **"AI POWERD CHATBOT"**
Of the Year 2020-21 towards the partial fulfillment of TY(MCA) as per SavitriBai Phule Pune
University.

Date:

Place:

Prof. Trupti U. Ahirrao
Project Guide

Prof. P.A.Chaudhari
H.O.D
Dept.of MCA(Engg)

Dr.P.C.Kulkarni
Principal

ACKNOWLEDGEMENT

Every work is the source that requires support from many people and areas. It gives us the proud privilege to complete the Project on **AI POWERED CHATBOT** under the valuable guidance and encouragement of our guide **Prof. TRUPTI U. AHIRRAO.**

I am also extremely grateful to our respected H.O.D. (MCA Eng. Dept.) **Prof.P.A.Chaudhari** and **Prof. Pranjali Deshmukh** (Project Coordinator) for providing all facilities and every help for the smooth progress of a Project.

I would also like to thank all the Staff Members of the MCA Department for timely help and inspiration for the completion of the Project.

"I would like to place on record our sincere thank to my friends for their helping during project development.

<div align="right">

Harshal Vinayak Shinde
(Roll no.22)
T.Y.MCA

</div>

ABSTRACT

A chatbot is a web application that is used to interact between a computer and a human in a natural language like humans chat. Chatbot chats with the user in a conversation in place of a human and reply to the user.

The goal of this report on chatbot was to resemble a human being in the way they interact, trying to make the user think he is chatting with another human being. The chatbot application helps the students to access university-related information from anywhere with an internet connection.

This system reduces the work of college administration providing information to students and also reduces the workload on the staff to answer all the queries of the students.

CONTENTS

Contents

List of Figures

1. INTRODUCTION

- Sometimes it is not possible to go manually to college for a small inquiry related to the department.

- Department will also be facing the big challenge that is how to answer the small queries of students.

- So Artificial intelligence is the solution to reduce the extra efforts of students and departments.

- Our main goal with this proposed system is a web application that gives a reply to the questions of the users.

- The user will get the appropriate answers to their queries. The answers will be given using artificial intelligence algorithms.

Problem definition

To design and develop an Artificial intelligence-based Chatbot for the website of the MCA department.

Objectives and Scope

Objective

Following are the objectives:

- Reduce management effort.

- Provide necessary details to students and parents online.

- This application enables students to be updated with college cultural activities.

- This application saves time for the student as well as teaching and non-teaching staff.

- Availability is 24x7

Scope

The scope of the system is limited to:

- Highly recommendable for college/Schools website.

- Highly recommendable for an online shopping website.

Organization of Report

Literature Survey:

Literature Survey includes the study of existing systems available with their problems and conclusion. A detailed description of the existing system which is manual also includes in the literature survey.

Software Requirement Specification

Software requirement specification includes software, hardware, and test data requirements.

Design and Modeling

Design and modeling include architectural design that contains all modules. Also include standard UML diagrams such as Class, Use case, and Sequence diagram.

Test Specification

Test specification contains a description of the test environment including test plan, testing tools, test cases, and test results.

Result and Analysis

Organize and present result data in the form of a table, chart, or graph.

2. Literature Survey

Existing system

- The existing system was manual and need some effort. It was a challenging task for answering each student's query manually. So the proposed system needs to come into focus.

- Many programs and applications have been developed based on this concept. But most of the common failure is also the poor processing which is not able to filter results in the time that can annoy people.

- Since then the problem is continuously studied using different conditions and we have come up with a new purposed system.

- The proposed system will simply take the query of the user which can be a student or a parent and will give a response according to the query.

- There are some existing systems as follows:

    (1) Eliza: Eliza makes the use of primitive natural language processing. It operates on the user's responses to scripts. ELIZA uses simple pattern matching techniques.

    (2) Doctor Script: DOCTOR is a simulation of a psychotherapist. It provided a basic human-like interaction with almost no information about human thought or emotion.

    (3) Cleverbot: Cleverbot is also a web application that converses with humans and uses artificial intelligence algorithms. It learns from human input.

    Drawbacks of the existing system

- ELIZA has a habit of parsing the responses of the user to frame the next question.

- Also after the conversation is continued these responses become predictable.

- DOCTOR did was that it used a lot of data from the statements from the humans to compose the responses by using simple templates.

- Cleverbot is constantly learning. Its data size is also increasing.

- Due to this, it appears to display a degree of "intelligence"

Proposed system

- In the proposed system we are going to see a model for an AI-Powered chatbot for the college website.

- The main goal of this proposed system is a web application that gives a reply to the questions of the users.

- Artificial intelligence will be used to answer the user's queries.

- The user will get the appropriate answers to their queries. The answers will be given using artificial intelligence algorithms.

- Users won't have to go personally to the college for inquiry.

- Chatbot makes use of machine learning to reach artificial intelligence helping them to understand the user query and provide an appropriate response.

- The chatbots are developed using the Artificial Intelligence Mark-up Language for communicating or interacting with the user.

The proposed system has the following objectives

- Reduce management effort.
- Provide necessary details to students and parents online.
- This application enables students to be updated with college cultural activities.
- Chatbot application saves time for the student as well as teaching and non-teaching staff.
- 24x7 availability for all students and parent's queries.
- This application saves time for the student as well as teaching and non-teaching staff.

The proposed system has the following Scope

- Highly recommendable for college/Schools website.
- Highly recommendable for an online shopping website.

3. Software Requirement Specification

The work of the system can be carried out using the following configurations-

Technology used

- Python3 , JQuery, HTML5, CSS3
- Flask, chatterbot corpus, sqlite3

Software requirements

- Operating System: Windows / Linux / macOS
- Web Browser, Python, Flask packages
- For college, server-side software's are required

Hardware requirements

- Internet connectivity
- Intel i3/AMD Ryzen3 processor-based computer
- 4 GB RAM, 320GB HDD

4. Design and Modeling

Architectural Design

Design is a multi-step process that focuses on data structure software architecture, pro- caesural details, (algorithms, etc.), and the interface between modules. The design process also translates the requirements into the presentation of software that can be accessed for quality before coding begins.

INPUT DESIGN

Input design is the process of converting user-originated inputs to a computer-based format. In the project, the input design is a chat-box design that takes input from a user as a message format.

OUTPUT DESIGN

Output design generally refers to the results and information that are generated by the system for many end-users; output is the main reason for developing the system and the basis on which they evaluate the usefulness of the application. In the project, the user will get the output as a response to the user's input.

Query Analysis and Response System:

This application provides information regarding the proceedings of the college. It is known that the answer will be given to the user if the input matches a pattern in the AIML files. In a case where the input does not match with the data in AIML files, keywords are fetched.
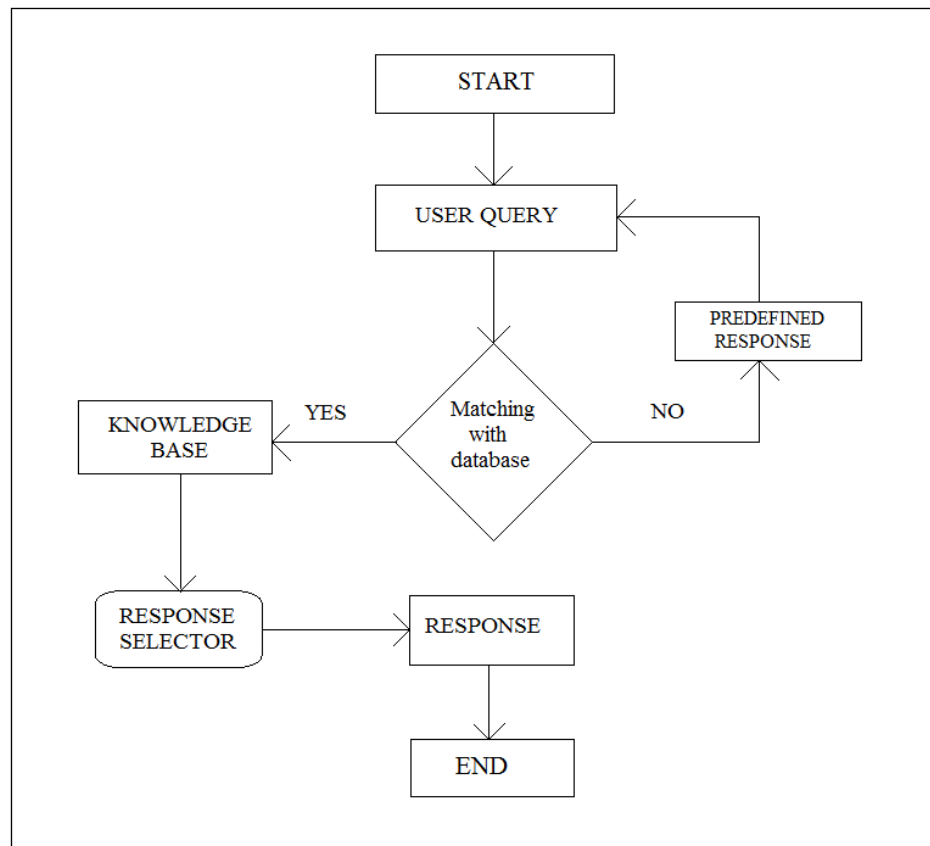
Flow Chart for the proposed system:

Figure 1: Flow Chart for the proposed system

AI Algorithm which is used in the proposed system:
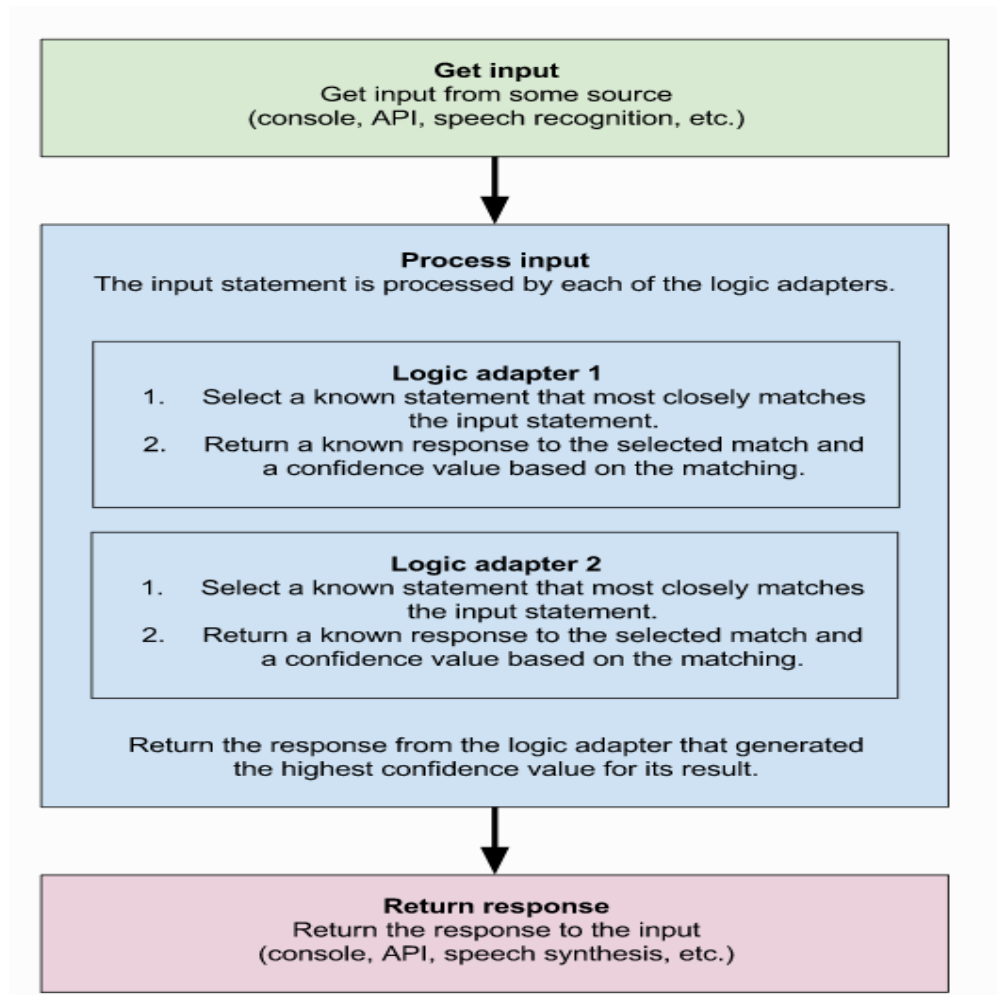
Figure 2: AI Algorithm which is used in the proposed system

College Related Response Activity:

Input Text → Response present in AIML files?

Yes → Send appropriate response

No → Find similarity in sense of input with available question set

Similar query found with confidence > 0.5

Yes → Send AIML response of the matched query

No → Store the input text in a log file for chatbot improvement by admin

Send a random response suggesting "Answer not available"

Figure 3: College-Related Response Activity

5. Implementation

A. AIML:

The question and answer pairs are stored in the AIML files so that the base for the normal conversations can be created. The patterns in AIML are matched with the input from the user and the appropriate answer is given.
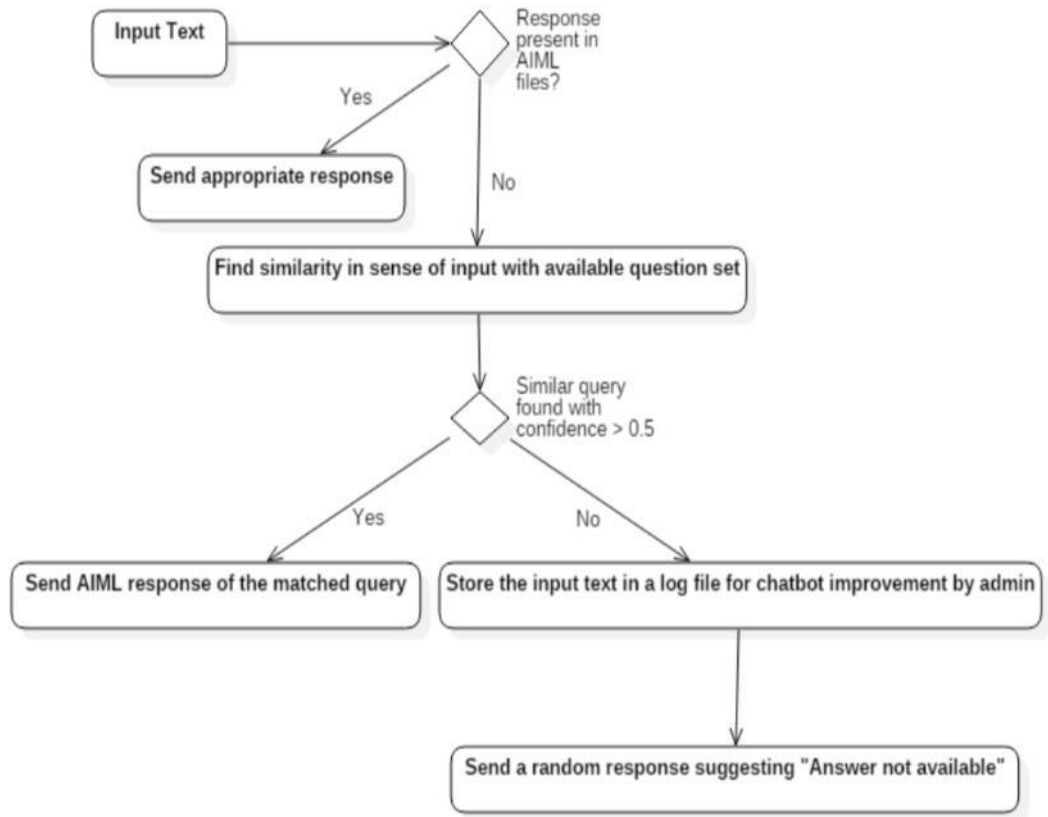
The sample AIML file structure is as follows:

```
<aiml version = "1.0.1" encoding = "UTF-8"?>
<category>
<pattern> HELLO USERNAME</pattern>
<template> Hello User! </template>
</category>
</aiml>
```

B. Lemmatization and POS Tagging Using WordNet:

Keywords were extracted from the input text which is known as information extraction. E.g. "What is the distance from Nashik?" contains "distance", "Nashik" as the keywords. Lemmatization and POS tagging were used to find appropriate Lemmas so that proper grouping can be done. Also, WordNet from Python 3's NLTK package was used for mapping the words to their base. E.g. walking, walk, and walked should map to walk.

C. Semantic Sentence Similarity

There are various cases in which the user might write the same sentence differently. E.g.,

Q1: Who is the HOD of the MCA department?
Q2: Tell me about the HOD of the MCA department.

Here both the questions have the same meaning. There can be many other combinations for the same query. The problem here is that all those combinations cannot be added to the file. Hence, the performance of the system will get affected. To solve this problem, the similarity is found between the input of the user and the data present in the question set. The input with the maximum score gets selected and then the proper answer is provided.

The above-mentioned score is calculated by averaging the similarity of the individual keywords of those sentences. The word with maximum similarity is matched between the keywords of the sentences. Then the similarity score is averaged to make the sentence similarity.

Another approach is the implementation of word similarity using the Path Similarity and Wu-Palmer (WUP) Similarity is used.

Like WordNet, a hierarchical structure is assumed where the path similarity compiles the shortest number of edges from one-word sense to another. It is to be noted that the longer word sense is less similar than those with a smaller path distance. E.g. car, bus, and cycle(we expect that a bus is similar to a car and not a cycle). The Wu-Palmer metric calculates the distance using the weights in the edges.

D. Log File:

A log file is also maintained which store those queries which the chatbot was not able to answer. It is the work of the admin to view the log file and the responses of the sentences. All this is done into the knowledge base. This is a very important method to improve the overall functioning of the chatbot

6. Testing

System Testing

System testing involves testing of software and successful running of the developed proposed system.

System testing is the stage of implementation that is aimed at ensuring that the system works accurately and efficiently before live operation commences. Testing is vital to the success of the system. System testing makes the logical assumption that if all the parts of the system are correct, then the goal will be successfully achieved. A series of testing is done for the proposed system before the system is ready for user acceptance testing.

The following are the Testing plans:

- Unit Testing

- Integration Testing

- Validation Testing

- Verification Testing

- User Acceptance Testing

Unit Testing

Unit testing focuses verification efforts on the smallest unit of the software design, the module. This is also known as module testing.

In this software, each module is tested. This testing was carried out during the programming stage. In this testing, each module is found to be working satisfactorily as regards the expected output from the module.

Integration Testing

One module can have adverse efforts on another. Integration testing is the systematic testing for the construction of program structure. Thus in the integration testing step, all the errors uncovered are corrected for the next testing steps.

Validation Testing

After integration testing, the software is completely assembled as a package, Invalidation testing functional or performance characteristics of an integrated package are checked.

Verification Testing

Verification is a fundamental concept in software design. This is the bridge between customer requirements and an implementation that satisfies those requirements.

User Acceptance Testing

The system under study is tested for user acceptance by constantly keeping in touch with the prospective system users at any time of developing and making changes whenever required.

• Test Cases

Test no	Test Description	Expected Result	Obtained Result	Status
01	Create a connection for the user to connect with a server.	Establish a connection between servers and generate a link.	The connection is established and a link is generated.	
02	Open a web browser and open a link	the page is open without getting a "404 Error"	The webpage is successfully opened without any error	
03	Ask a Query related to the MCA department	Getting a response to a query asked to MCAbot	The response is received of a query related to the MCA department	
04	Ask a Query that is not related to college or MCA department	Getting a response to a query asked to MCAbot	The response is received of a query that is not related to the MCA department	