ALEC - (Artificial Linguistic Enquiry Chatbot)

A Mini Project Report Submitted
in Partial Fulfillment of the Requirements
for the Degree of
Master of Computer Applications

by
Aishwarya Sadana(2017CA12)
Aditya Bhawsar(2017CA59)
Mansi Sharma(2017CA79)
Pavan Chandravanshi(2017CA56)

under the supervision of
Dr. Anoj Kumar
Associate Professor, CSED

to the
COMPUTER SCIENCE & ENGINEERING DEPARTMENT
MOTILAL NEHRU NATIONAL INSTITUTE OF TECHNOLOGY
ALLAHABAD, PRAYAGRAJ, U.P.
May, 2019

UNDERTAKING

We declare that the work presented in this report titled "ALEC-*Artificial Linguistic Enquiry Chatbot*", submitted to the Computer Science and Engineering Department, Motilal Nehru National Institute of Technology, Allahabad, for the award of the *Master of Computer Applications* degree, is our original work. We have not plagiarized or submitted the award of any other degree. In case this undertaking is found incorrect, we accept that our degree may be unconditionally withdrawn.

May, 2019
Prayagraj

Aishwarya Sadana(2017CA12)

Aditya Bhawsar(2017CA59)

Mansi Sharma(2017CA79)

Pavan Chandravanshi(2017CA56)

CERTIFICATE

Certified that the work contained in the report titled **"ALEC-*Artificial Lirguistic Enquiry Chatbot*"**, by **Aishwarya Sadana(2017CA12), Aditya Bhawsar (2017CA59), Mansi Sharma(2017CA79) and Pavan Chandravanshi(2017CA56)**,has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

Dr. Anoj Kumar
Associate Professor, Computer Science & Engineering Department
M.N.N.I.T Allahabad, Prayagraj

May, 2019

Preface

A **Chatbot** is a computer program or an artificial intelligence which conducts a conversation via auditory or textual methods.A Chatbot for any website helps to find out the queries through chat. We can get the answers instantly without searching it in the whole website and if the query is not available it asks to submit the query later the admin will add the answers for it. It is an advanced and promising expression of interaction batween human aad machine. ALEC that we made, using PHP, HTML, CSS, JavaScript and Python, caters all of these requirements and helps the students to find answers to their queries instantly.

Acknowledgements

It is obvious that the development of this project needed the support of many people. We would like to express my deepest appreciation to all those who provided us the possibility to complete this project. First of all, we would like to thank our Head of Department, **Prof. A. K. Singh**, to give us the opportunity to take up this project.

Special thanks goes to our project mentor,**Dr. Anoj Kumar**, under whose guidance we have developed this project. We are indebted to him for his constant support and encouragement to bring this project in its present form. With his able guidance and suggestions only we could develop ALEC.

Finally, we would also like to raise a vote of thanks to our classmates who have supported and motivated us in every possible way they could. Their reviews mattered a lot during the development of this project.

Contents

Preface **iv**

Acknowledgements **v**

1 Introduction **1**

1.1 Problem Statement 1

1.2 Objective 2

1.3 About The Project 2

1.4 Motivation 3

1.5 Related Work 3

2 Design Methodology Used **4**

3 Proposed Work **5**

3.1 Functional and Non-Functional Requirements 5

3.1.1 Functional Requirements 5

3.1.2 Non-Functional Requirements 6

4 Design **7**

4.1 Data Flow Diagrams 7

4.1.1 Level 0 7

4.1.2 Level 1 7

4.1.3 Level 2 8

4.2 Use Case Diagram 8

5 Implementation **9**
5.1 Module Details 9
5.1.1 Admin 9
5.1.2 User 10
5.1.3 Notice Board 10
5.2 Corpus Data sets 10
5.3 Python Socket Server 10

6 Technology and Tools Used **11**
6.1 Python 11
6.2 PHP 11
6.3 HTML and CSS 11
6.4 JavaScript 12
6.5 Natural Language ToolKit (NLTK) 12
6.6 ChatterBot Library 12
6.7 MySQL with phpmyadmin 13
6.8 XAMPP 13

7 Software/Hardware Requirements **14**
7.1 Software Requirements 14
7.2 System Requirements 14

8 Testing **15**
8.1 Functional Tests 15

9 Experimental Results **16**

10 Conclusion and Future Work **21**

11 References **22**

Chapter 1

Introduction

A.L.E.C. stands for "Artificial Linguistic Enquiry Chatbot". It will act as a one place solution to all the queries asked by the students. It will help the students to get the desired information at the right time. It will be like a virtual assistant to which we can ask questions and get instant answers.

1.1 Problem Statement

The website of Computer Science and engineering Department contains all the information required by students. But it is very difficult to search for the information on the site without any prior knowledge of where it could be found. Students need to enquire information about the department, courses, subjects, past projects, etc. They need to search through the website to look for answers. It can be time consuming or sometime information is not present on the website. Students need to manually visit to the college to get their queries answered by the college help desk. This process consumes lot of time as well as money as the customer needed to visit college if its miles away from home. Also, this process may lead to communication gap between student and college.

1.2 Objective

To create a Chat Bot for the Computer Science and Engineering Department which will help student's to get answer to their query, immediately and efficiently.

1.3 About The Project

Chat bots typically provide a text-based user interface, allowing the user to type commands and receive text as well as text to speech response. Chat bots are usually stateful services, remembering previous commands in order to provide functionality. When chat bot technology is integrated with popular web services it can be utilized securely by an even larger audience.

A.L.E.C. is built with python's **ChatterBot** library using machine learning algorithms that analyzes user's queries, understand user's message and provide right answer. This System will be a web application which provides answer to the query of the student very effectively. Students just have to put their query to the bot which is used for chatting. The system will use the artificial intelligence algorithms to give appropriate answers to the user. The student will not have to go to the college for enquiring something. Student can use the chat bot to get the answers to their queries at any point of time. This system may help students to stay updated with the college activities.

This is an advanced PHP,HTML, CSS, JavaStript and Python Chatbot system that will allow you to chat with the machine and get a detailed information in return. It includes:

- Chat view.
- Active communication method
- Notice Board for CSED

1.4 Motivation

In this fast growing technical world being student of a technical institution we still have to find answer to our query manually by going to the dapartment, which makes us old school and can cost us our time and money too. We want to develop a web application which can give us answers to our queries of any type at any time in minimum cost. We sometimes pass our time by chatting with different ChatBot available on internet, so to make one of them was indeed an interesting idea. For learning aspect, we wanted to work on a full web-application and apply all the concepts they we have learned in the past years.

1.5 Related Work

Since currently we do not have any Chatbot for CSED, we had to look for other sources for related work. **Cleverbot[2], Disha[3], etc.** are some frequently used chatbots but none of them can fulfill our requirement.

Chapter 2

Design Methodology Used

We used the basic Waterfall Model to Develop our project.
In waterfall model we follow the following steps to develop the project

1. Gather information
2. Design
3. Implement
4. Perform testing
5. Maintenance

This model is very simple to understand and implement. However, due to its simplicity, this model is only useful for a limited type of projects.

Following are the reasons why we choose this model:

1. Our requirements were rigidly defined from the very beginning.
2. The technology that we were going to use was very well defined.
3. Each phase of the project was clearly defined.

Chapter 3

Proposed Work

The system that we are planning to build should support the following basic functionality:

- Provides user friendly chat view.
- Provides reference links.
- Provides tucked information.
- Live notice board and announcements.
- Admin can manage unanswered questions.
- User can submit the queries.

3.1 Functional and Non-Functional Requirements

3.1.1 Functional Requirements

A.L.E.C must support the following features:

- The admin can manage unanswered queries and can add it in database if required.
- The chatbot should be responsive to users.

- Chatbot extracts keyword from the query and provide answers based on it.
- Only authorized users can upload notices to notice board.

3.1.2 Non-Functional Requirements

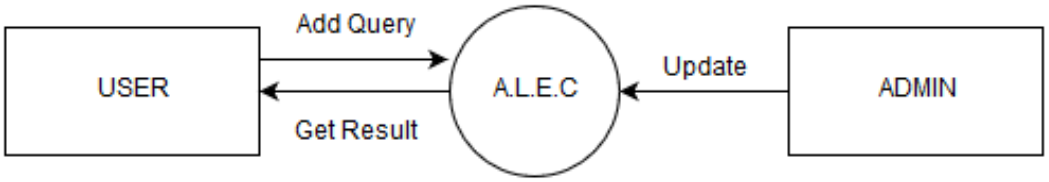
- The chatbot should be designed in such a way that it should be easy for others to maintain and add new features.
- The chatbot should be visually appealing.
- The web-app should support web page caching to reduce the response time.
- All the functionality of the web-site should be easily testable using automated Unit Tests.

Chapter 4

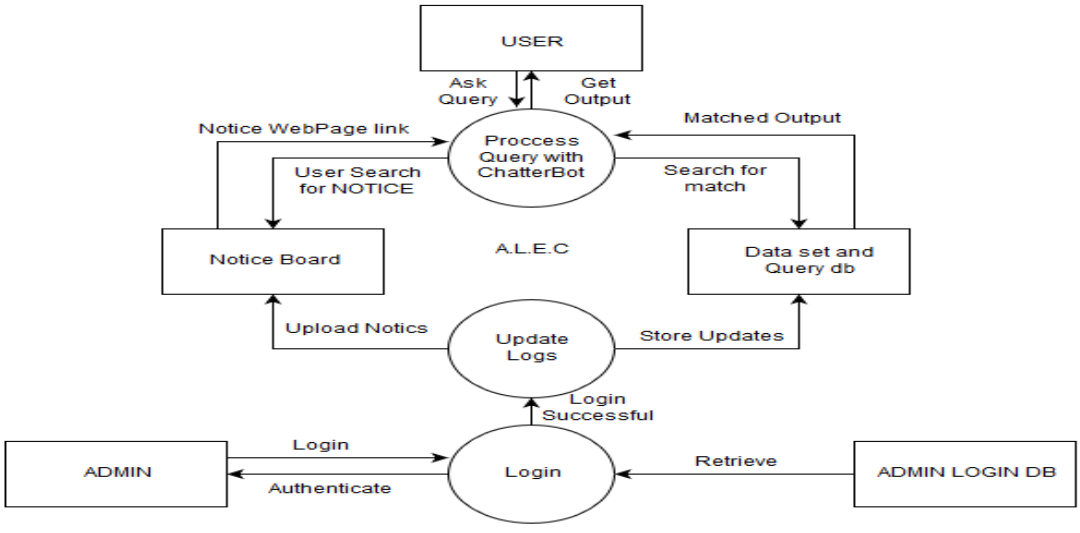
Design

4.1 Data Flow Diagrams

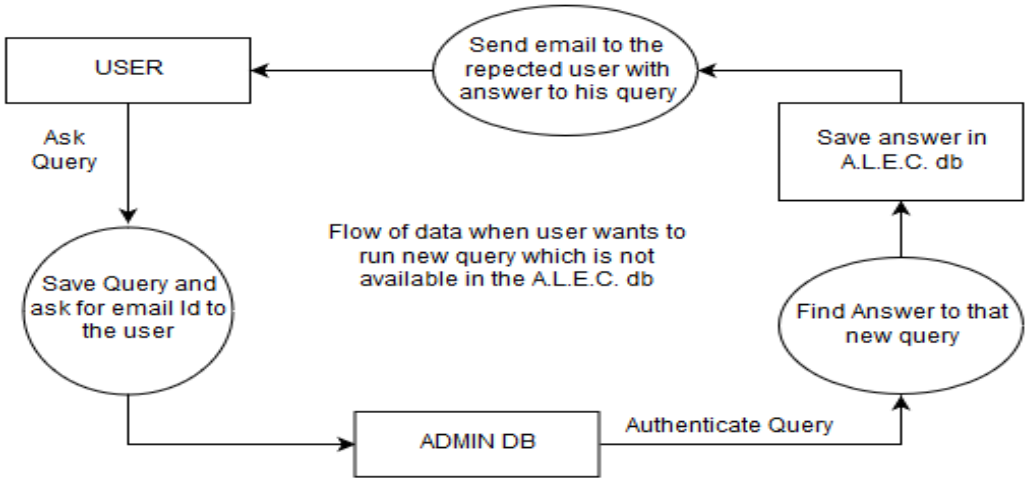
4.1.1 Level 0



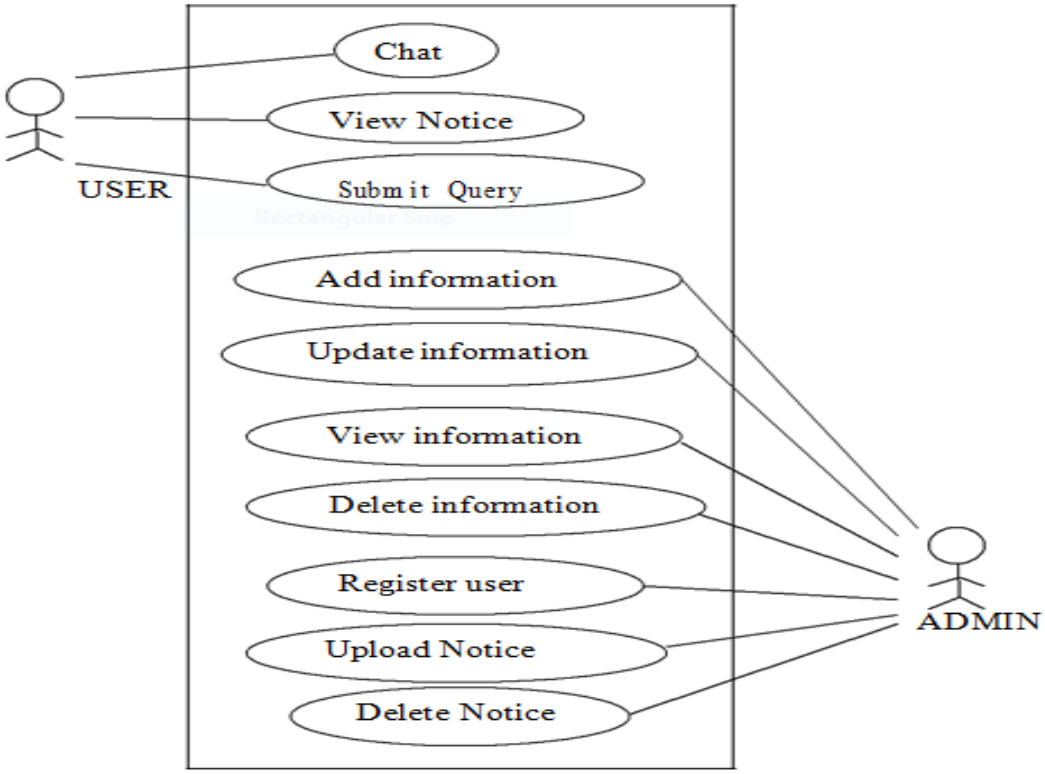
4.1.2 Level 1



4.1.3 Level 2



4.2 Use Case Diagram



Chapter 5

Implementation

5.1 Module Details

A Chatbot have limited number of modules due to its simple and effective functions. We have divided A.L.E.C. into three modules admin, user and notice board. These modules are created simultaneously to cater to the needs of students of CSED.

5.1.1 Admin

- View information
- Update information
- Add information
- Delete information
- Register user to upload notice
- Upload notice
- Delete notice

5.1.2 User

- Ask Query
- Submit Query
- View Notice

5.1.3 Notice Board

This module is defined for displaying the important notices issued by the department. Students can view the notice board by just typing Notice in the chatbot.

5.2 Corpus Data sets

Corpus is a large collection of texts. It is a body of written or spoken material upon which a linguistic analysis is based. A.L.E.C uses .yml(extension) text files for data sets. Like,
– – Good morning! How are you doing?
–I am doing very well, thank you for asking.
– –Youre welcome.
–Do you like hats?

5.3 Python Socket Server

SimpleWebSocketServer, this asynchronous WebSocket handler create connection with server and client and send data frame to the client. This library have various functions to facilitate user with packet transferring. If data is a unicode object then the frame is sent as Text. If the data is abytearray object then the frame is sent as Binary.

Chapter 6

Technology and Tools Used

6.1 Python

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting of glue language to connect existing components together. Large set of available library makes it useful for ALEC.

6.2 PHP

PHP (recursive acronym for PHP: Hypertext Preprocessor) is a widely-used open source general-purpose scripting language that is especially suited for web developmmnt and can be embedded into HTML.

6.3 HTML and CSS

HTML(Hypertext Markup Language) and CSS(Cascading Style Sheets) are two of the core technologies for building Web pages. HTML provides the structure of the page, CSS the (visual and aural)layout, for a variety of devices. Along with graphics and scripting, HTML and CSS are the basis of building Web pages and

Web Applications.

6.4 JavaScript

JavaScript often abbreviated as "JS", is a high-level, dynamic, untyped, and interpreted run-time language. In has been standardized in the ECMA Script language specification. Alongside HTML and CSS, JavaScript is one of the three core technologies of World Wide Web content production; the majority of websites employ it, and all modern Web browsers support it without the need for plugins. JavaScript is prototype-based with first-class functions, making it multi-paradigm language, supporting object-oriented, imperative, and functional programming styles. It has an API for working with text, arrays, dates and regular expressions, but does not include any I/O, such as networking, storage, or graphics facilities, relying for these upon the host environment in which it is embedded.

6.5 Natural Language ToolKit (NLTK)

The Natural Language Toolkit (NLTK) is a platform used for building Python programs that work with human language data for applying statistical natural language processing (NLP). It contains text processing libraries for tokenization, parsing, classification, stemming, tagging and semantic reasoning. It also includes graphical demonstrations and sample data sets as well as accompanied by a cook book and a book which explains the principles behind the underlying language processing tasks that NLTK supports.

6.6 ChatterBot Library

ChatterBot is a Python library that makes it easy to generate automated responses to a user's input. ChatterBot uses a selection of machine learning algorithms to produce different types of responses. This makes it easy for developers to create chat bots and automate conversations with users.

6.7 MySQL with phpmyadmin

MySQL is an Oracle-backed open source relational database management system(RDBMS) based on Structured Query Language (SQL). phpMyAdmin is a free software tool written in PHP,intended to handle the administration of MySQL over the Web. phpMyAdmin supports a wide range of operations on MySQL.

6.8 XAMPP

XAMPP is a free and open-source cross-platform web server solution stackpackage developed by Apache Friends, consisting mainly of the Apache HTTP server, MariaDB database and interpreters for scripts written in the PHP. Since most actual web server deployment use the same component as XAMPP, it makes transitioning from a local test server to a live server possible.

Chapter 7

Software/Hardware Requirements

7.1 Software Requirements

- Python
- Xampp Server

7.2 System Requirements

- Intel Core (or Dual Core 2GHz) or equal AMD CPU.
- 2GB RAM or Above
- 40GB of Free HDD Space

Chapter 8

Testing

8.1 Functional Tests

Since not everything gets covered in Unit Tests, we performed manual functional tests ta make sure that the basic functionalities do not have a bug. Some of the features that we have tested are:

- Accessing all the information on the site (Answers, Results and more).
- Multiple tabs working simultaneously without affecting each other.
- Adding/Deleting notice for authorised person and admin.
- Viewing the Web-App on multiple size windows so that we are sure it is completely responsive.
- Trying to access modification rights without Admin Access.

Chapter 9

Experimental Results

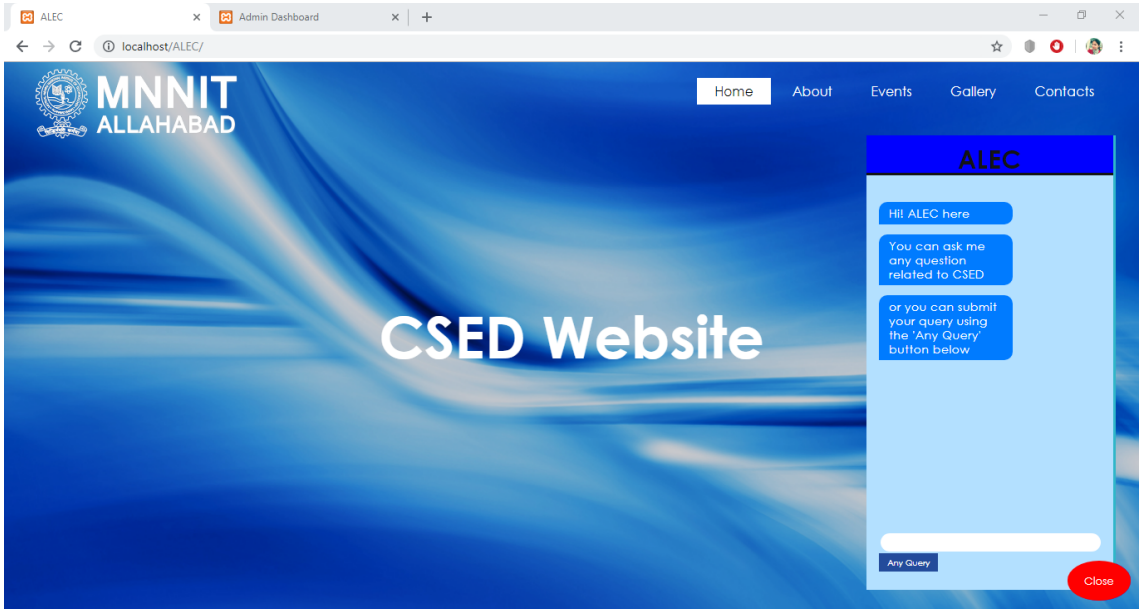


Fig 9.1: *Initial view of chatbot*

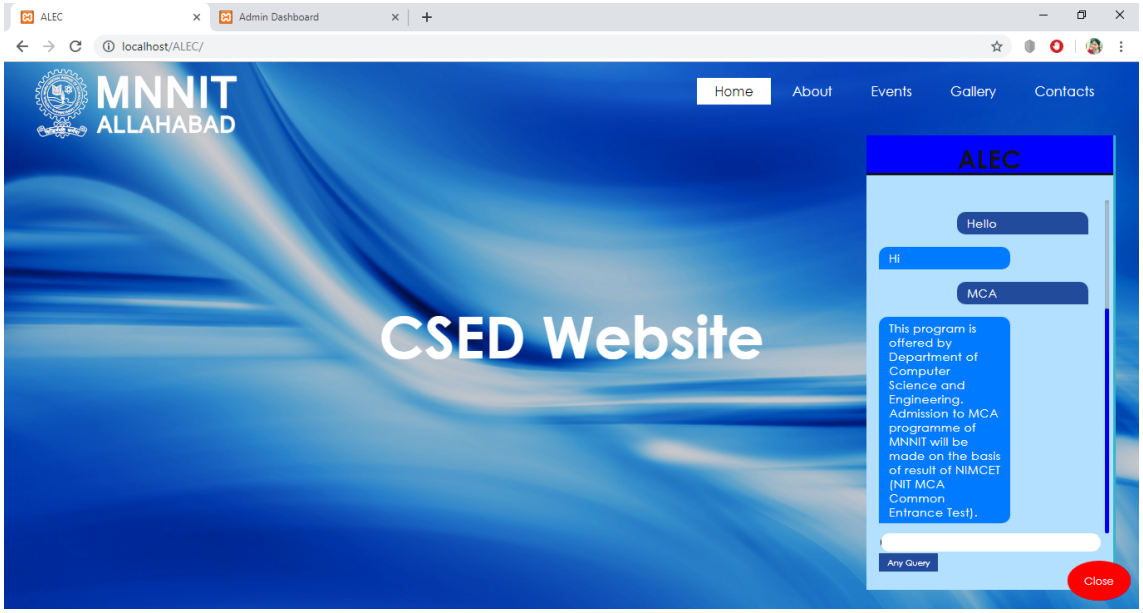


Fig 9.2: *Reply for MCA given by chatbot*

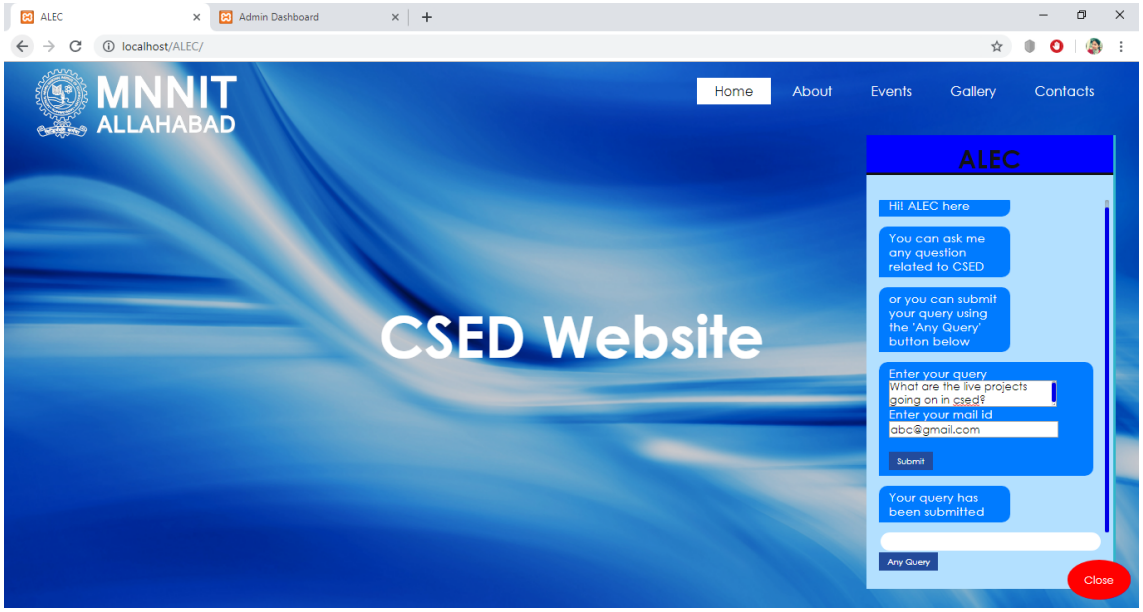


Fig 9.3: *User submitting query*

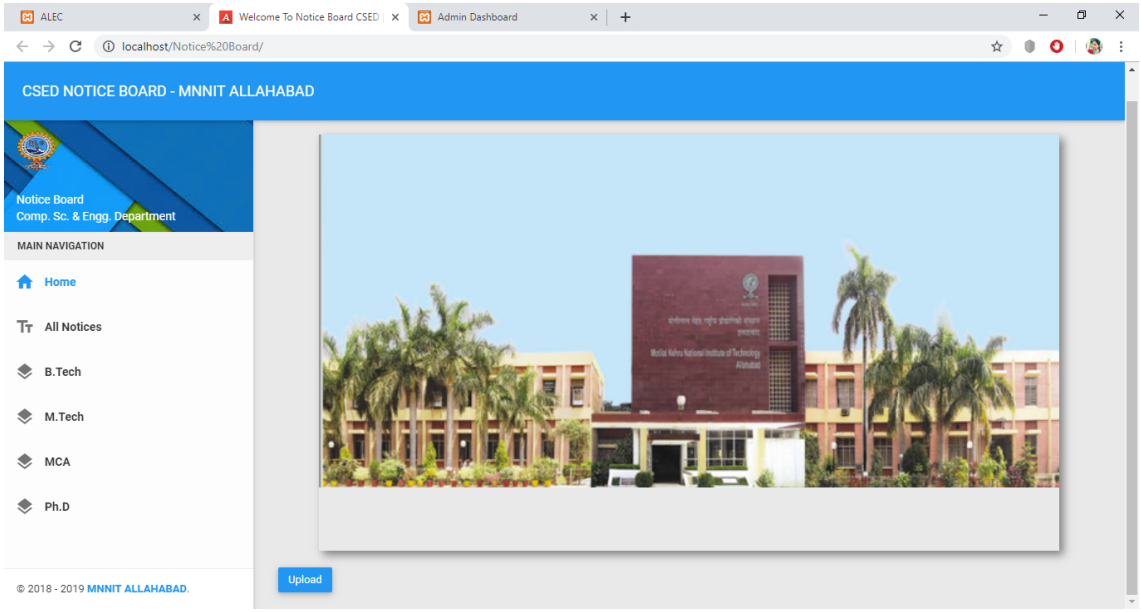


Fig 9.4: *Notice Board*

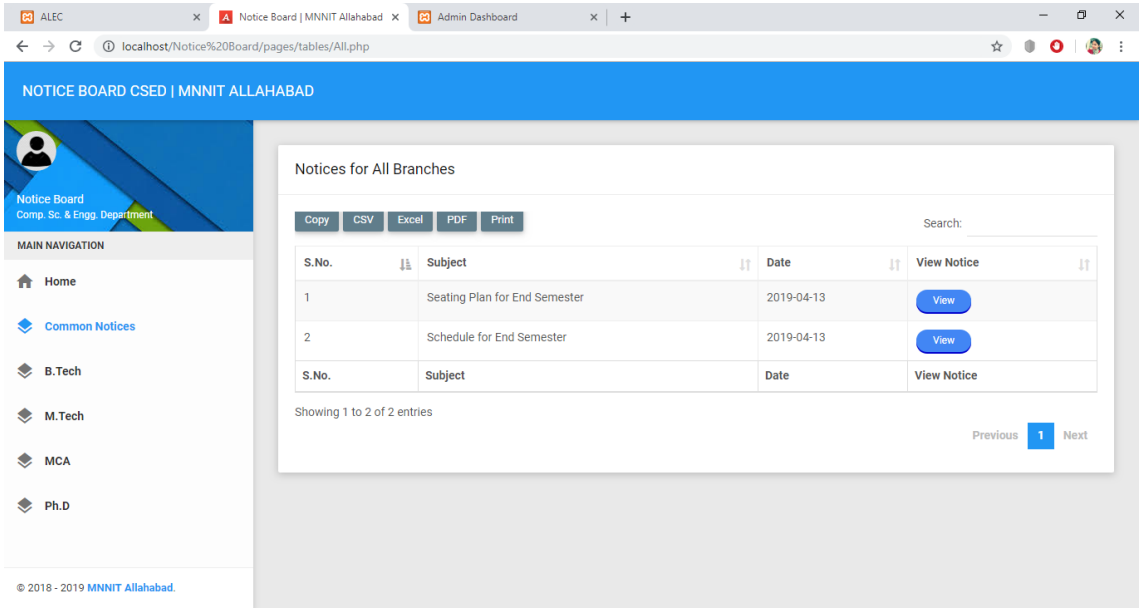


Fig 9.5: *Notices for all Branches*

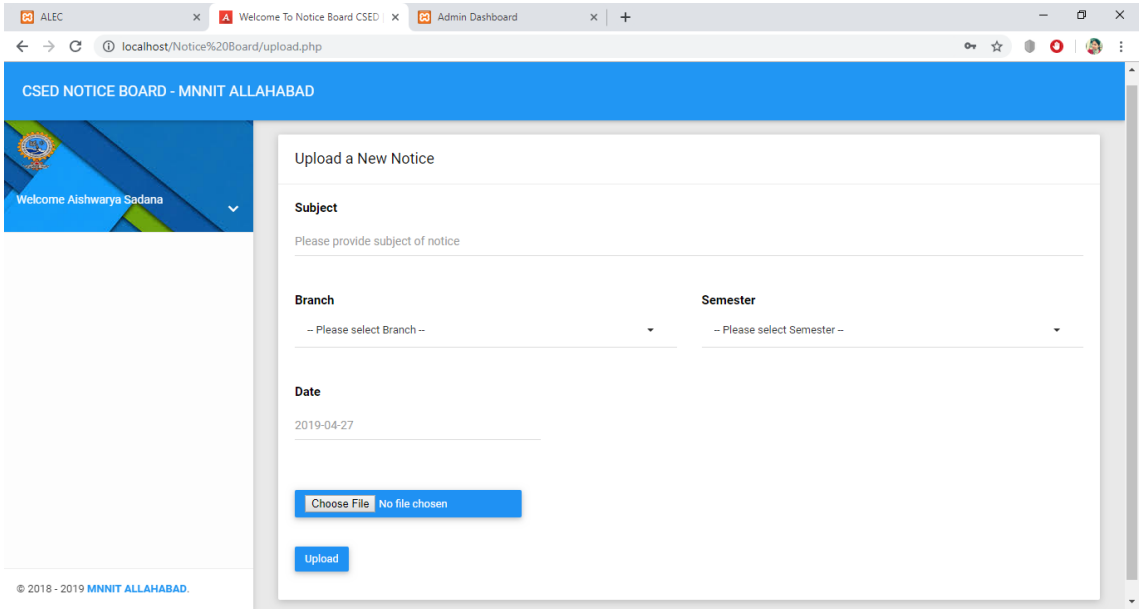


Fig 9.6: *Notice Board - Upload notice*

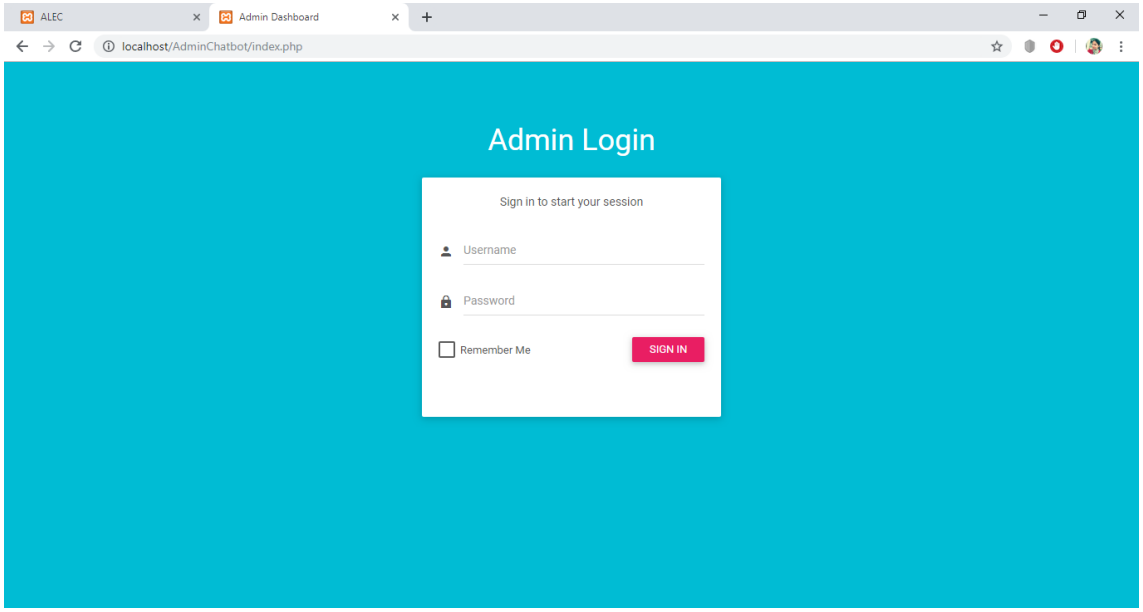


Fig 9.7: *Admin log in*

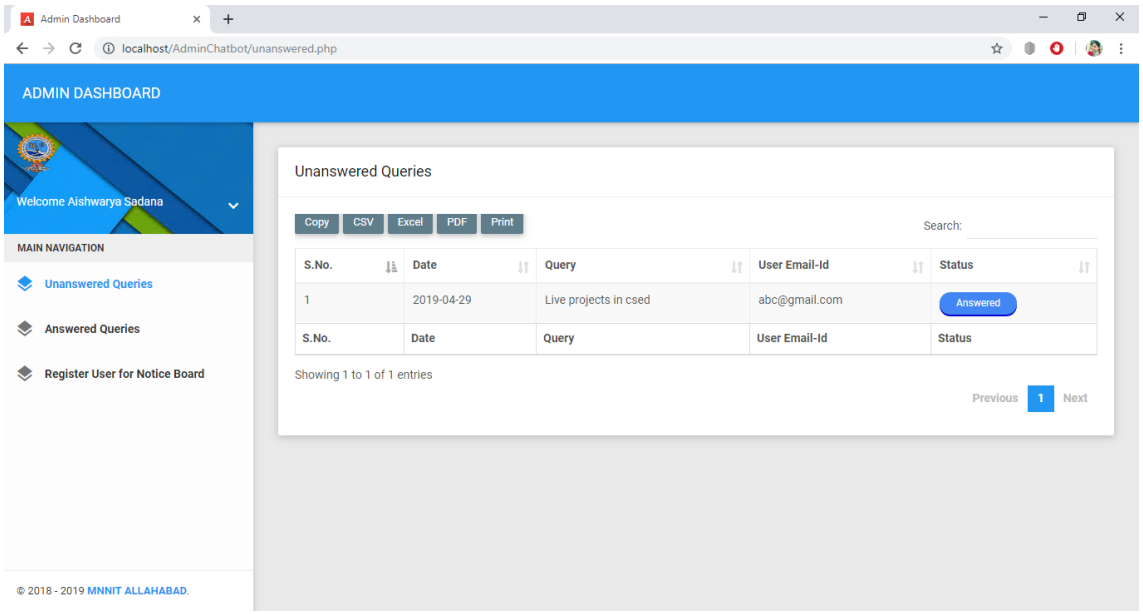


Fig 9.8: *Admin Dashboard - Unanswered Queries*

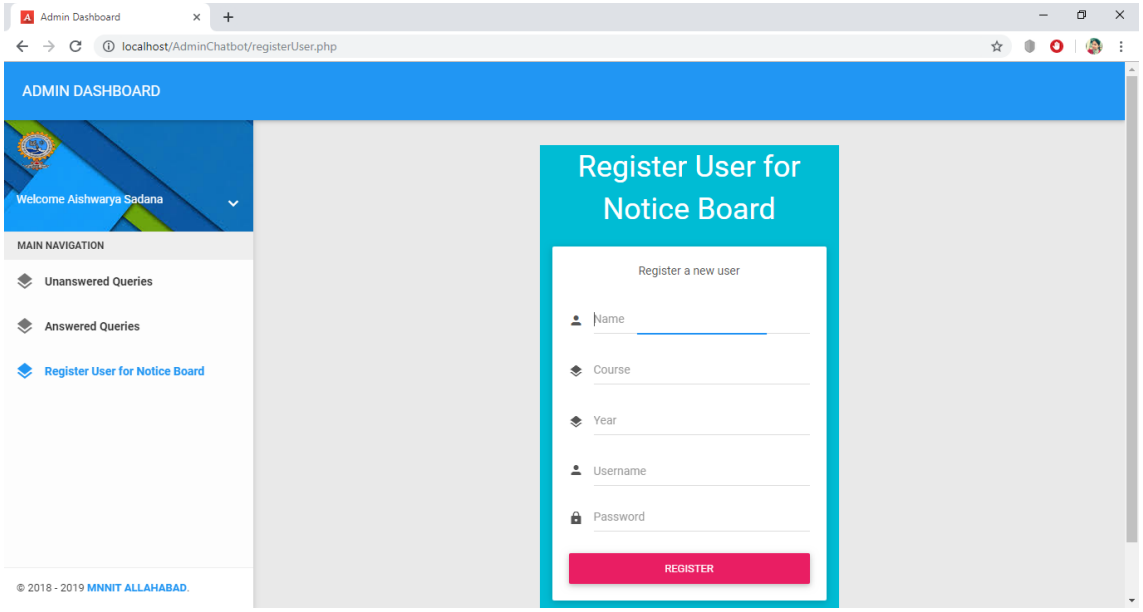


Fig 9.9: *Registration for notice board*

Chapter 10

Conclusion and Future Work

We now have a Web-app for our CSED Department that will be useful for the students to find answers to their queries. Though it may have a very limited number of features as of now, it can very easily be modified to add new features.

In future many other functionalities can be added to the portal like:

- If the answer is not proper or invalid, then as per user suggestion it can be modified by the admin
- Answers with the available options so that the user does not have to type his response.(predicted answers will be available to choose)
- Image and PDF in the answers
- Speech to text and text to speech feature.
- More effective notice board view.
- User login for more specific queries like assignment, marks or exam information.
- Publishing to social media platform.

Chapter 11

References

1. https://drive.google.com/file/d/0B-tCvLzyt01FYlQ2dVRBWEtTNkE/view
   (Research Paper referred)

2. https://www.cleverbot.com/
   (Normal user friendly chatbot)

3. https://www.irctc.co.in/nget/
   (A railway guide on IRCTC website)

4. https://en.m.wikipedia.org/wiki/Xampp
   (Xampp server)

5. https://chatterbot.readthedocs.io/en/stable/
   (Python Chatterbot library)