AI CHAT BOT

PROJECT REPORT

*Submitted by*

Lasata Tuladhar (5-2-410-12-2014)

Rabina Sthapit (5-2-410-25-2014)

Rabince Shrestha (5-2-410-26-2014)

*in partial fulfillment for the award of the degree*

*Of*

BACHELOR OF SCIENCE COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

IN

PRIME COLLEGE, KATHMANDU

(TU AFFILIATION)



AI CHATBOT

PROJECT REPORT

*Submitted by*

Lasata Tuladhar (5-2-410-12-2014)

Rabina Sthapit (5-2-410-25-2014)

Rabince Shrestha (5-2-410-26-2014)

*in partial fulfillment for the award of the degree*

*of*

BACHELOR OF SCIENCE COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

IN

PRIME COLLEGE, KATHMANDU



TRIBHUVAN UNIVERSITY

June, 2018

ACKNOWLEDGEMENT

We owe a debt of gratitude to Mrs.Dikshya Singh - our coordinator and our supervisor for the vision and foresight that inspired us to complete this project. Also express our special thanks to Mr. Krishna Shrestha – our principal for providing us the platform to learn.

We would also like to thank our parents for the constant support and their inspirations. We are highly indebted to Prime College and their staffs for constantly allowing us to use the equipment’s and all the help we needed.

Abstract

This project is related with building a chatbot by using artificial intelligence system. We chose this specific topic by looking at the present context of our world being taken over by AI. Humans are always fascinated for making their daily life easier with new technologies. Chat bots are among those many technologies that are growing at a very high pace. The main objectives of this project is to allow people to interact with the bots whenever they have some queries. This will reduce the work load of companies by allowing bots to handle the customer queries. The result of this project will save people their time for having to wait for a response. With a little advancement in the knowledge base, it will answer the queries more effectively by making it a lot like two humans interacting. As the reasoning power of machine is defined by knowledge base, it has the efficiency of pattern matching algorithm. Here English language is used as communication language and the mode of communication is written form. It is developed using python programming and its intelligence is based on the capacity of the knowledge base. Here we try our best to include as inclusive knowledge as we can.

Keywords: Intelligence, knowledge base, pattern matching, machine.

**Table of Contents**

[**CHAPTER I** 1](#_Toc514432773)

[1. INTRODUCTION 1](#_Toc514432774)

[1.1 Background 1](#_Toc514432775)

[1.2 Problem Definition 2](#_Toc514432776)

[1.3 Objectives 2](#_Toc514432777)

[1.4 Scope and Limitations 2](#_Toc514432778)

[1.4.1 Scope 2](#_Toc514432779)

[1.4.2 Limitations 3](#_Toc514432780)

[**CHAPTER II** 4](#_Toc514432781)

[2. REQUIREMENT ANALYSIS AND FEASIBILITY ANALYSIS 4](#_Toc514432782)

[2.1 Literature Review 4](#_Toc514432783)

[2.2 Data Collection Method 4](#_Toc514432784)

[2.2.1 Sources of Data 4](#_Toc514432785)

[2.3 Requirement Specification 5](#_Toc514432786)

[2.3.2 Non-Functional Requirements 6](#_Toc514432787)

[2.4 Feasibility Analysis 6](#_Toc514432788)

[2.4.1 Operational Feasibility 6](#_Toc514432789)

[2.4.2 Technical feasibility 7](#_Toc514432790)

[2.4.3 Economic Feasibility 7](#_Toc514432791)

[2.4.4 Schedule Feasibility 7](#_Toc514432792)

[2.5 Structuring System Requirements 9](#_Toc514432793)

[2.5.1 Data Modeling 9](#_Toc514432794)

[2.5.2 Process Modelling 9](#_Toc514432795)

[**CHAPTER III** 11](#_Toc514432796)

[3. SYSTEM DESIGN 11](#_Toc514432797)

[3.1 System Architecture and Overview 11](#_Toc514432798)

[3.2 System Diagram 12](#_Toc514432799)

[**CHAPTER IV** 16](#_Toc514432800)

[1. SYSTEM IMPLEMENTATION AND TESTING 16](#_Toc514432801)

[4.1 Implementation Overview 16](#_Toc514432802)

[4.2 Tools Used 18](#_Toc514432803)

[4.2.1 Front End Tools 18](#_Toc514432804)

[4.2.2 Back End Tools 18](#_Toc514432805)

**List of Figures**

[Fig 1. ER Diagram 10](#_Toc514264445)

[Fig 2. Context Diagram 11](#_Toc514264446)

[Fig 3. DFD Level-1 Diagram 12](#_Toc514264447)

[Fig 4. System Architecture of Chat bot 13](#_Toc514264448)

[Fig 5. System Diagram 14](#_Toc514264449)

[Fig 6. Use Case Diagram 16](#_Toc514264450)

[Fig 7. Activity Diagram 17](#_Toc514264451)

[Fig 8. Agile Method 19](#_Toc514264452)

**List of Tables**

[Table 1 User-End Requirement 6](#_Toc514340300)

[Table 2 System-End Requirement 6](#_Toc514340301)

[Table 3. Grant Chart 8](#_Toc514340302)

# **CHAPTER I**

# 1. INTRODUCTION

## 1.1 Background

Artificial Intelligence is the broader concept of machines being able to carry out tasks in a way that we would consider “smart”. It attempts not just to understand but also to build intelligent entities. Whereas Machine Learning is a current application of AI based around the idea that we should really just be able to give machines access to data and let them learn for themselves. As technology and importantly, our understanding of how our minds work, has progressed, our concept of what constitutes AI has changed. Rather than increasingly complex calculations, work in the field of AI concentrated on mimicking human decision making processes and carrying out tasks in ever more human ways. Humans are always fascinated with making their daily lives easier with the implementation of new technology. Chatbots are one such means of technology which helps humans in a lot of ways. As chatbots can be applied to any service industry, humans who want to do no matter what will be making use of chatbots to make it easier.

A chatbot is a service that people interact with via a chat interface. The main concept is to ask questions using voice or by typing in the same way we would ask a person. The chatbot will usually respond in a conversational style, and it may carry out actions in response to our conversation. It often runs inside a popular messaging application, such as Facebook Messenger, Slack, or SMS. It answers our question, rather than directing us to a website.

Chatbots are basically categorized into two main types:

1. Rule based

They have their own kind of knowledge base. It responds to specific commands by simply checking its knowledge database. If the user doesn’t use the right commands or sets of action, the bot may not be able to provide us with the answers we desire.

1. Natural language

It has the ability to constantly learn from user interactions to become better at predicting their needs. It can chat in a similar way like humans.

The machine learns and uses its algorithms through supervised and unsupervised learning. Supervised learning means to train the machine to translate the input data into a desired output value. Unsupervised learning means discovering new patterns in the data without any prior information and training.

## 1.2 Problem Definition

Several chatbot applications have been developed till date in the field of artificially software development. They have their own knowledge base and feature but most of them are too straight forward in application. Whenever it gets input from user, it simply checks its knowledge database and answers it. The approach we have used here can be quite similar to the existing applications, but we have made certain advancements in the knowledge base. The pattern matching is made more efficient and the answers to the user question is also made more like that of humans.

## 1.3 Objectives

The major goal of initiating the project is to impart some intelligence power to a software so that it could communicate with a user. The major objectives include:

* To implement the concept of machine learning and its application in a computer software
* To implement Natural language processing
* To provide automated customer services
* To reduce operational and service expense

## 1.4 Scope and Limitations

### 1.4.1 Scope

It is intended to cover several knowledge areas of English language in this software. Topics like introduction, general knowledge, universe, feelings, personalities etc. have been included in this application. Availability of Chatbots 24/7 with the immense knowledge they can hold is all set to outperform humans. With the speed and accuracy, they are offering support to enterprises, they will soon augment human capabilities. Users love to interact with Chatbots as it saves them time and in most of the cases offers them clear and concrete answers. For the best result, there must be large set of training data proper pattern matching algorithm, and realistic answers.

### 1.4.2 Limitations

Since chatbots already use Natural Language Processing and machine learning, they can handle whole customer support conversations without any hindrance. Let us take into consideration a case where customer has an unusual request or needs individual handling. Chatbots cannot fully replace humans, but they should complement each other. So, in this case, an agent should interact with a customer in second instance. This ensures uninterrupted conversation with the customer resulting in higher customer satisfaction.

# **CHAPTER II**

# 2. REQUIREMENT ANALYSIS AND FEASIBILITY ANALYSIS

## 2.1 Literature Review

Evolution of artificial intelligence was the most important thing for the development of chatbot. Alan Turing’s proposal suggested that machine can also behave like human beings. From 1950 to now there has been a lot of improvement and development in artificial intelligence.

## 2.2 Data Collection Method

The various data needed for our project were collected from different sources. It includes both the primary data and the secondary data.

### 2.2.1 Sources of Data

* Internet research

The internet is a compelling tool for research. It enables efficient, cost-effective data collection and facilitates access to large samples of data. Internet played a vital role during our project. We came to know about various chatbot applications that has been introduced so far. Ongoing through various reports regarding the chatbot, we gained the knowledge about further betterment of our developing applications.

* Interaction

We interacted with our supervisor about what features can be added to our chatbot. We consulted our seniors on the subject matter and collected as much information we could.

* Existing reports

Like the voice recognition system used by virtual assistant like Google Now, Apple’s Siri, and Microsoft’s Cortana, there are various chat-based applications too. Among them Replica, messenger bots are some few examples where we learned how the bots interact with the user in the same way as human. This helped us to learn kind of information is fed and how they keep learning from the previous examples.

## 2.3 Requirement Specification

1. User-End Requirements

Table 1 User-End Requirement

|  |  |
| --- | --- |
| Rid | Requirements |
| 1. | User should enter text in the textbox to chat. |
| 2. | User should enter proper formatting of sentences. |
| 3. | User should enter correct spelling of words. |
| 4. | User should maintain proper arrangement of words |
| 5. | User should type meaningful sentences. |
| 6. | User should type in English font. |
| 7. | User may or may not use the punctuation marks. |
| 8. | User should maintain arrangement of noun, pronoun or verb. |
| 9. | User must form the sentences at least having one noun, pronoun or verbs. |
| 10. | User should not enter text more than 50 characters. |
| 11. | User should not take more than five minutes of time to reply. |

1. System-End Requirements

Table 2 System-End Requirement

|  |  |
| --- | --- |
| 1.. | Chatbot should take any form of input under English language. |
| 2. | Chatbot should accept sentences with at least one noun, pronoun or verb. |
| 3. | Chatbot should not take more than 5 seconds to reply. |
| 4. | Chatbot should give meaningful answer. |
| 5. | Chatbot should entertain the user. |
| 6. | Chatbot should accept only correct spelled words. |
| 7. | Chatbot should maintain history of chat. |
| 8. | Chatbot should give answer related to topic. |
| 9. | Chatbot should not repeat the same answer for the same question. |
| 10. | Chatbot may tell the same thing in different way. |
| 11. | Chatbot should have knowledge base related to different topics. |
| 12. | Chatbot should give default answer to the questions it does not know. |
| 13. | Chatbot should reply in a polite way. |
| 14. | Chatbot should give the most informative answer as a reply. |

### 2.3.2 Non-Functional Requirements

1. Performance requirement:

* The risk factor was analyzed at initial step for better performance of the application.
* The overall performance of the application is reliable and enables the user to use the application efficiently.

1. Security requirements:

* The application should be well secured from the non- authoritative action of modifying knowledge base.

1. Flexibility:

* The application should flexible as new domain can de added.

1. Maintainability

* If any error occurs, it should be easily maintainable.

1. Scalability:

* It should be scalable based on addition of knowledge base.

1. Testability:

* It should be testable for various forms of testing.

## 2.4 Feasibility Analysis

It is used to determine the viability of an idea with an emphasis on identifying the potential problems. It attempts to answer one main question, “Will the idea work and should be proceed with it?” Feasibility is the measure of how beneficial or practical an information system will be to an organization.

### 2.4.1 Operational Feasibility

The main focus is on providing customer services to the general public by answering their queries. This will help organization to balance the work load of the employees. On the other hand, customers do not have to wait for long to get a reply regarding their questions. In addition to connecting the user to a specific person when the bots do not have the exact answers for their queries.

### 2.4.2 Technical feasibility

It also measures the availability of technical resources and expertise. For our AI chatbot projects we are targeting for desktop applications. We will utilize existing technology that will lower the risk of our project. We will be using python for programming as it is compatible for all OS.

### 2.4.3 Economic Feasibility

When specific requirements and solutions have been identified we can weigh the cost and benefits of the alternatives. The hardware and software used are simple and no additional hardware requirement. It is based on the existing system, so the cost will be minimum. The only cost that will be encountered is the printing and electricity expenses.

### Schedule Feasibility

This includes our project schedule and the time we allocated for their completion.

The grant is as follows:

Table 3. Grant Chart

|  |  |  |  |
| --- | --- | --- | --- |
| Task | Start Date | Due Date | Duration(days) |
| Requirement Collection | 15-03-18 | 20-03-18 | 5 |
| Planning | 21-03-18 | 31-03-18 | 10 |
| Analysis | 01-04-18 | 10-04-18 | 10 |
| Design | 11-04-18 | 22-04-18 | 11 |
| Implementation | 22-04-18 | 26-05-18 | 36 |
| Testing | 27-05-18 | 02-06-18 | 7 |
| Documentation | 03-06-18 | 10-06-18 | 7 |

## 2.5 Structuring System Requirements

### 2.5.1 Data Modeling

Fig 1. ER Diagram

### 2.5.2 Process Modelling

User

No

Request

1.0

Check Login

Response

Yes

Chatbot

Admin

Verify Response

Fig 2. Context Diagram

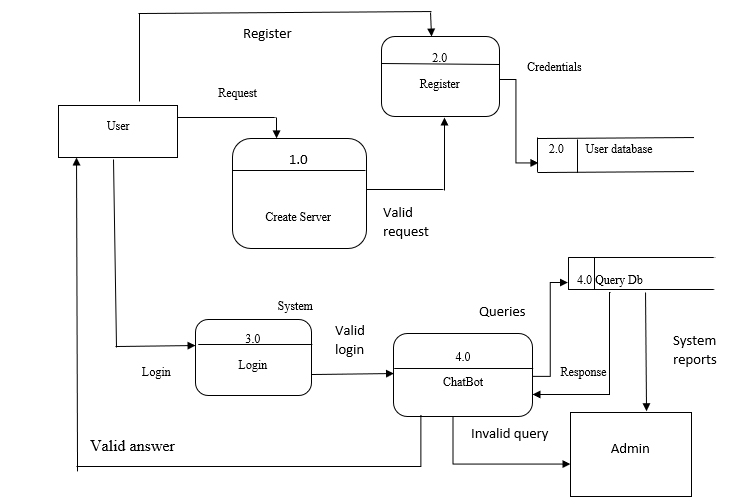


Fig 3. DFD Level-1 Diagram

# **CHAPTER III**

# 3. SYSTEM DESIGN

## 3.1 System Architecture and Overview

The concept is to type the text by the users and the bot analyzes and responds to the question by looking through its knowledge base. The bot learns the actions based on the training data provided and also from the previous examples.

Program

Knowledge Base

Dialog saved in log file

Choice of response

Comparison of query with the rules

Chat Bot

Query Response

User

Fig 4. System Architecture of Chat bot

## 3.2 System Diagram

Input text Understanding

Searching Knowledge Base

Output Generation



Fig 5. System Diagram

The block diagram shown above is described below:

1. Text Input:

The input to the system is some text written by the user through user interface. As it is based on English language, the font has to be English there. User us supposed to provide the appropriate ordering of the words, grammar and the proper symbols as necessary.

1. Conversion:

The process of conversion involves converting the inputted word segment into appropriate forms. It involve conversion to lower case and removal of the unnecessary symbols like apostrophes (‘), question mark (?), and full stop (.).

1. Segmentation:

As user types certain things in the UI, the text needs to be segmented. The text is segmented based on the use of space. If there are n number of spaces, there are n+1 number of words in the input sentences.

1. Knowledge Base:

The knowledge base is the central component of the chatbot. It stores knowledge in the form facts and rules. Usually predicate logic is used for this purpose. The knowledge base may also store the confidence level that a fact or rule is true or valid. Here, the knowledge base contains both declarative knowledge (facts about objects, events and situation). Although many knowledge representation techniques have been used in such system, the most prevalent form of knowledge representation currently used in expert system is the rule-based productions system approach.

1. Pattern Matching:

Here, patterns are already saved in the forms= to text files of text files in the knowledge base. It first searches all the knowledge base and find the matched pattern on the basis of matched words. The pattern with highest number of matched words is selected as for the reply. If two or more than two patterns are matched, then we calculate the output value is selected. Again, if output value is same for both patterns, then any one of them is selected randomly. The pattern matching algorithm is also explained under the section pattern matching algorithm is also explained under the selection pattern matching algorithm of system theory.

1. Output Selection:

From among the large list of the available knowledge base, the pattern matching algorithm finds the best one. For each of the questions, there are at least 3 sets of answers so selection is to be made out of these answers. These sets of answers are made to remove the repetition of answers during communication. Repetitive answers are boring in dual communications so three implementation having similar meanings are kept in the knowledge base. Since three of them have similar meanings, any of them can be selected randomly.

1. Result Generation

Finally, the selected output is presented to the user in the user interface.

3.2.5. UML Class Diagram

User Question

Get User Input

Search Knowledge Base

Match

Satisfactory

Yes No

Get the Answer

Word Match

Yes No

Select one Answer

Collect questions with equal match

Reply

Calculate numerical value for questions

Select Weights

Default Answer

Save to new questions

Fig 7. Activity Diagram

# **CHAPTER IV**

# SYSTEM IMPLEMENTATION AND TESTING

## 4.1 Implementation Overview

There are various methodology in software development practices. Among them we chose the agile methodologies. Agile methods refer to the incremental and iterative method of management. It focuses on helping teams in an evolving landscape and maintaining a focus on the rapid delivery of business value. Agile methods are flexible, and they perfectly address the needs of customers. During the whole cycle, user involvement is encouraged, providing visibility & transparency, showing the actual progress of projects.

Benefits in agile methods includes:

* Customer satisfactions through updates and continuous delivery
* Flexible approach
* High quality products
* Reduced risk

Basic workflow in agile methods

Requirements

Test and Feedback

Architecture and Design

Implementation

Fig 8. Agile Method

* Requirements

This phase starts by identifying the necessary requirements based on the business strategy and objectives. All possible requirements of the system to be developed are captured and documented in a requirement specification document.

* Architecture and design

The requirement specifications from first phase are studied in this phase and the system design is prepared. It prepares a solution architecture that is needed before construction begins. It also establishes the standards for screen design, programming standards, data management, platform provisioning, user acceptance testing approach and more.

* Implementation

It is the framework for addressing these issues. With inputs from the system design, the system is first developed in small programs called units, which are integrated in the next phase. Each unit is developed and tested for its functionality, which is referred to as Unit Testing.

* Test and Feedback

All the units developed in the implementation phase are integrated into a system after testing of each unit. The entire system is tested for any faults and failures. Once the functional and non-functional testing is done, the product is deployed in the customer environment or released into the market for feedback.

## 4.2 Tools Used

### 4.2.1 Front End Tools

* Html/ CSS/ JavaScript : Sublime text
* Adobe Photoshop: For designing

### 4.2.2 Back End Tools

* Python: Django Framework
* PostgreSQL : Database