ABSTRACT:

The college faculties face a rough time searching for data in the cabinets or computers and as a result they take a long time before they could collect the data, do the analysis, interpret and then present the result. Also a possibility of erroneous presented data is common in manual analysis. The Computers store the data effectively but there arises a need for a mining tool for analysis and interpretation and a HTML-CSS front end to present the result interactively. This report provides the ways to implement an application to act as an artificially intelligent user-friendly agent to mine the available information feedback from students and faculties to present it appropriately. It will also perform the appropriate probability and numerical calculations for setting up question papers, assigning salary and many more. The user interaction is either through touch or voice input and the output is viewing or downloading the related information.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **CHAPTER** | **TITLE** | **PAGE NO.** |
|  | ABSTRACT | iv |
|  | LIST OF FIGURES | viii |
|  | LIST OF ABBREVIATIONS | ix |
| 1 | INTRODUCTION | 1 |
| 1.1 | AIM | 1 |
| 1.2 | SYNOPSIS | 1 |
| 2 | SYSTEM STUDY | 2 |
| 2.1 | LITERATURE SURVEY | 2 |
| 2.2 | EXISTING SYSTEM | 3 |
| 2.3 | PROPOSED SYSTEM | 3 |
| 3 | SYSTEM ANALYSIS AND REQUIREMENT SPECIFICATIONS | 4 |
| 3.1 | INTRODUCTION | 4 |
| 3.2 | REQUIREMENT SPECIFICATION | 5 |
| 3.2.1 | HARDWARE REQUIREMENTS | 5 |
| 3.2.2 | SOFTWARE REQUIREMENTS | 5 |
| 3.3 | TECHNOLOGIES USED | 5 |
| 3.3.1 | JAVA | 5 |
| 3.3.2 | DATA MINING | 8 |
| 3.3.3 | APACHE TOMCAT SERVER | 8 |
| 3.3.4 | MYSQL | 9 |
| 4 | SYSTEM DESIGN | 10 |
| 4.1 | SYSTEM ARCHITECTURE | 10 |
| 4.2 | USE CASE DIAGRAM | 11 |
| 4.3 | SEQUENCE DIAGRAM | 12 |
| 4.4 | CLASS DIAGRAM | 13 |
| 4.5 | ACTIVITY DIAGRAM | 14 |
| 4.6 | COLLABORATION DIAGRAM | 15 |
| 4.7 | DATAFLOW DIAGRAM | 15 |
| 5 | SYSTEM DESIGN | 17 |
| 5.1 | MODULES | 18 |
|  | APPENDIX A | 20 |
|  | APPENDIX B | 42 |
|  | REFERENCES | 81 |

**LIST OF ABBREVATIONS**

|  |  |
| --- | --- |
| HTTP | Hyper Text Transfer Protocol |
| IP | Internet Protocol. |
| JDK | Java Development Toolkit |
| JMF . | Java Media Framework |
| TCP | Transmission Control protocol |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **FIGURE NO** | **NAME OF THE FIGURE** | **PAGE NO** |
| 3.1 | Java Interpretation and Compilation | 2 |
| 3.3 | Java Platform | 13 |
| 4.1 | System Architecture | 17 |
| 4.2 | Usecase Diagram | 18 |
| 4.3 | Sequence Diagram | 19 |
| 4.4 | Class Diagram | 20 |
| 4.5 | Activity Diagram | 21 |
| 4.6 | Collaboration Diagram | 22 |
| 4.7 | Data Flow Diagram | 22 |
|  |  |  |

CHAPTER 1

INTRODUCTION

The 21st century has brought about many innovations which are mostly in automation to make the lives of people easier and better to live. People look up on technology to artificially sense their needs for reducing their burden and the time taken to complete their jobs. People expect the machine which is an artificial intelligent equipment to sense his / her needs. We would also want to relate relevant day to day information and feedback to get a different perspective to the available information which is useful for interpretation, analysis and evaluation. The artificially intelligent student information network would make all of the above a possibility under a single cloud.

We have already seen how the usual college management works. The attendance, the marks, the syllabus, the time table and all are drawn up or recorded manually. This actually takes some time and is also not perfect. We may not be the perfect judge as to which a tough subject is or who the best teacher to handle crisis is. Further students and faculties would want up to date information in their hands to decide on anything such as the study pattern, the important topics to be covered and all that. So we are bringing out an artificially intelligent application which heavily uses data mining to present the best knowledgeable information about the best teacher for each subject, the rating for the teacher and student likewise, performance of teachers and likewise for students, time table framing, identified tough subjects and question paper framing. The aforementioned information is presented dynamically and not by the administrator but by the Artificial Intelligent agent which utilizes Data Mining.

SYSTEM OVERVIEW

The Artificial Intelligent Agent for the application can be implemented using the Natural Language Processing for voice based input and set of other codes written for arithmetic and probability calculation using the Python Programming Language. In case of a voice input, there are 3 stages which is the syntactic, semantic and pragmatic analysis. This will basically act as a decision support system. For implementing the Data Warehouse for our application, we make use of MySQL and for analysis and interpretation we make use of Python Database API to connect to MySQL database. We utilize the Python Flask API to fetch the results from Python code and publish it in the server at Local host. The new JavaScript Web Speech API makes it easy to add speech recognition to your web pages. This API allows fine control and flexibility over the speech recognition capabilities in Chrome version 25 and later.

OBJECTIVE

Our objective is to store the data easily, analyze and interpret the results effectively and error free without the help of humans and present the result in an interactive and understandable manner requiring lesser or zero key strokes to be performed for giving input.

SYSTEM STUDY

The system study phase analyses the demerits of the existing system, defines the objectives to be achieved by the new solution and evaluates other solution alternatives.

LITERATURE SURVEY

An extensive study was made on different existing applications which inspired us to come up with this application for the college. Siri in Apple IOS allows you to use your voice to send messages, schedule meetings, place phone calls, and more. But the functionalities of Siri are not specific and in depth . For example, we can’t ask Siri to compute a salary waiver for each faculty or draw up time tables which we regularly do manually in college as a result. Enterprise Resource Tools such as Fedena Pro Support are in their native state and requires human skill to extract vital information from available patterns. The brain child of this innovation has to be the WebFOCUS Business Intelligence platform. It does higher end analytics and presents graphical comparative and vital information. But while this focuses only on analytics and interpretation, our application focuses on the higher level college activities such as periodic calculation of salary of teachers based on college performance and rating and also with the help of this application, we can set up question papers with different range of difficulties. The book on the Natural Language Processing with Python by Steven Bird, Ewan Klein and Edward Loper tells you how Python Programming language is simple and powerful for including excellent functionality for processing linguistic data and presenting it appropriately to the user. We can construct a powerful data analysis tool using MySQL and Mining tool using Python. Further the following papers were a guiding light in our research and development of our project.

Details: Y. C. Huei, "Benefits and introduction to python programming for freshmore students using inexpensive robots," Teaching, Assessment and Learning (TALE), 2014 International Conference on, Wellington, 2014, pp. 12-17. doi: 10.1109/TALE.2014.7062611

Abstract: Robots were used for teaching at various levels of education. Many research papers reported positive results of using robots to aid students in understanding of programming concepts and writing programs. In this paper, the investigation results had shown that mini-projects using the Finch robot not only enhanced creativity, problem solving skills, and research capability of students but also had helped understood Python programming concepts and writing programs in a collaborative teaching environment at Singapore University of Technology and Design (SUTD). The pedagogical activities of this research were described. Then the experiences and results of findings were shared. The shared knowledge will be beneficial to academics or educators who wish to integrate robots into their programming course, especially in a course of more than 250 students, students with numerous interests, and students with different learning abilities.

Details: V. Aswini and S. K. Lavanya, "Pattern discovery for text mining," Computation of Power, Energy, Information and Communication (ICCPEIC), 2014 International Conference on, Chennai, 2014, pp. 412-416. doi: 10.1109/ICCPEIC.2014.6915399

Abstract: Text mining can be defined as the art of extracting data from large amount of texts. It allows to structure and categorize the text contents which are initially non organized and heterogeneous. Text mining is an important data mining technique which includes the most successful technique to extract the effective patterns. The paper focuses on developing an efficient method for discovering patterns from the document. In text mining field, pattern mining techniques are used to find text patterns, such as frequent item sets, closed frequent item sets, co-occurring terms. This paper presents an innovative and effective pattern discovery technique which includes the process of pattern evolving and pattern deploying, to improve the effectiveness of using and updating discovered patterns for finding relevant and interesting information.

EXISTING SYSTEM

The existing system does hold a database of related college information wherein the user can view the stored information and can do further analysis. Further users with administrative privileges will be responsible to store the data in the database.

DISADVANTAGES

1. It mostly presents you non analyzed and non interpreted raw data.

2. Information has to be viewed in the database with no proper basic User Interface.

3. It is strenuous to add the data in the database in the existing system.

4. More keyboard input and multiple screens involved to view the required and necessary output.

PROPOSED SYSTEM

In the proposed system, a python script handles faster insertion, updating and deletion of data for the administrator requiring limited amount of inputs to be entered. Further it is an easy aid for the administrator to add the proper data. Viewing the populated and the interpreted data is much simpler. It just requires two screens: one to handle input through voice and the other to present the analyzed and interpreted output data in an interactive and user friendly manner without the help of the administrator artificially.

ADVANTAGES:

1.Faster: Python makes the overall system much faster than the older systems.

2. Artificial Interpretation: The application can artificially interpret what the user speaks and saves the time taken in manual selection of commands.

3.Simple UI: Lesser number of screens thus making it a simple Interface for the user to understand.

4.Automated Data mining: Automatic analysis and pattern extraction done entirely by the python script with the information a voice input away.

SYSTEM ANALYSIS

Systems are created to solve problems. One can think of the system approach in an organized way of dealing with a problem. The analysis phase is the second phase of SDLC and to decide if the existing project should go ahead and can be implemented with the resources available. This also includes looking at the existing system to see what it is doing for the organization and how well the system is doing its job. The feasibility of the project is also considered and the group has to ask questions such as

• Can this system be created with the resources we have?

• Will this system significantly improve the organization?

• Does the old system necessarily need to be replaced?

PROBLEM STATEMENT

2.1.1 Problem Definition

1. Analyzed and interpreted information are more representative.

2. No dedicated tool and UI for detailed college system information storage and data extraction.

3. The systems that are currently in use lack quality and accuracy.

2.1.2 Problem Identification

This project is mainly focused on the efficient use of college records to easily establish a relationship between different entities such as Student, teacher, marks secured and performance to determine multiple information such as department pass percentage, tough subject, teacher salary and many more. There are many colleges who do not have dedicated people to simply handle and play with the department data and so our project does away with the need for such a person.

2.1.3 Problem Analysis

The most common solution to the problem would be to assign a proper person to work with the college records. It could not be feasible in all the cases. Also, it is very time consuming and inefficient. So the most probable solution would be an online environment to collect the college records, analyze it and present it interactively on its own. Further we utilize the Natural Language Processing algorithm to get the user input, get the required input and do the necessary required actions. The results are presented based upon a closer understanding of voice input and the collected set of related and identified keywords from the list. Then the Flask API effectively renders the output online through the server in an Interactive manner.

2.2 SCOPE

2.2.1 Product Scope

The product scope is about the product itself. The product scope defines what the product will look like, how will it work, its features, etc. Our system will play a major role in improving the quality of the college system and it could be available for users throughout the college and with the help of professionals, the accuracy of the system could be improved and made available for use easily. The product has a wider range of scope in other domains as well.

2.2.2 Business Scope

Business scope contains Textual Scope Description, list of external actors and their roles, description of business activities. The actors of the system are the user who uses the system and the server which hosts the system itself. The users could be the students or teachers or even administrators who needs suggestions regarding their disease. The server on the other hand, accesses all the data and helps the retrieval of the same regarding the users and use those data to fetch the needed college information.

2.3 PURPOSE

The purpose of the project is to provide a standalone server-system to handle the entire college information and to make a seriously capable artificial Intelligent System to understand what the user says and provide him with what he needs.

2.4 RESOURCES REQUIRED

Resources are required to carry out the project tasks. They can be people, equipment, facilities, funding required for the completion of a project activity. Here the Artificial application is the information provider and the user or the sometimes even the administrator is the knowledge seeker. The development environment required should have a minimum configuration of i3 core processor. The software resources include the operation system, Python interpreter and other editors.

2.5 FEASIBILITY STUDY The feasibility study is to determine whether the solution is achievable, given the organizations, resources and constrains .By performing feasibility study the scope of the system will be defined completely. Most computer systems are developed to satisfy a known user requirement. Once the decision is made a report is forwarded and is known as feasibility report. Feasibility study is an evaluation of system regarding to its workability, impact on the organizations, ability to meet user needs and effective use of resources. The existing system is pasteurized above and based upon the identification of pitfalls in the existing system and relative inconvenience and difficulties, an outline for a new system is conceptualized and framed. During feasibility analysis, the following primary areas were considered very carefully in our project. The feasibility studies are under 3 contexts

i. Technical Feasibility

ii. Economic Feasibility

iii. Operational Feasibility

2.5.1 Technical Feasibility

What resources are available for the given developer system?

Is the problem worth solving?

In the proposed system technical feasibility centers around the hardware and the software and to what extent it can support the proposed system. The tools are used to develop the application are the best tools available in the technological scenario and hence it requires efficient and versatile programmers and programming skills. Even though the technical requirements are needed for development of our project, any system which has a web browser and internet connectivity could easily use our product. Hence the proposed system is technically feasible.

2.5.2 Economic Feasibility

Economic Feasibility is used for evaluating the effectiveness of the system. The procedure is used for determining the cost and the benefits or the savings that are expected from the system and compare with the cost. In our project all the software obtained are open source and so no cost is incurred apart from awarding the developer for his effort in putting all his programming skills to design such a system. So for sure the proposed system is cost effective than the existing system. Hence the proposed system is economically feasible.

2.5.3 Operational Feasibility

The main problem in the developing the new system is getting acceptance from the college to provide their details to the system. Another problem is cooperation from the users because many users are reluctant to operate new systems. The product being developed is hosted on a webserver so the users are spared from the tedious job of installing software systems. The developed system is usable by range of people who have basic knowledge to surf the internet. Further the admin is given a basic Python UI to add, update or delete record details. Hence the system is operationally feasible.

14 2.6 SYSTEM REQUIREMENTS

2.6.1 Hardware Requirements Processor :

• i3 core processor

• Google Chrome with version 25 and above

Active Internet Connection

2.6.2 Software Requirements

Platform : Windows XP / Ubuntu

Front End : Python Flask

Back End: MySQL

2.7 OTHER REQUIREMENTS

2.7.1 Functional Requirements Functional requirements may be calculations, technical details, data manipulation and processing and other specific functionality that define what a system is supposed to accomplish. Behavioral requirements describing all the cases where the system uses the functional requirements are captured in use cases. The plan for implementing functional requirements is detailed in the system design. A function is described as a set of inputs, the behavior, and outputs.

2.7.2 Non Functional Requirements A non-functional requirement is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors. The plan for implementing non-functional requirements is detailed in the system architecture.

*TECHNOLOGIES USED*

PYTHON

A useful data type built into Python is the dictionary. Unlike sequences, which are indexed by a range of numbers, dictionaries are indexed by keys, which can be any immutable type; strings and numbers can always be keys. Tuples can be used as keys if they contain only strings, numbers, or tuples; if a tuple contains any mutable object either directly or indirectly, it cannot be used as a key. A pair of braces creates an empty dictionary: {}. Placing a comma-separated list of key: value pairs within the braces adds initial key: value pairs to the dictionary; this is also the way dictionaries are written on output. The main operations on a dictionary are storing a value with some key and extracting the value given the key. It is also possible to delete a key: value pair with DEL. If you store using a key that is already in use, the old value associated with that key is forgotten. It is an error to extract a value using a non-existent key. The keys() method of a dictionary object returns a list of all the keys used in the dictionary, in arbitrary order (if you want it sorted, just apply the sorted() function to it). To check whether a single key is in the dictionary, use the “in” keyword.

Flask library is a light weight web framework in Python. Flask is fun and easy to set up. Flask is considered more [Pythonic](http://blog.startifact.com/posts/older/what-is-pythonic.html) than Django because Flask web application code is in most cases more explicit. Flask is easy to get started with as a beginner because there is little boilerplate code for getting a simple app up and running.

The DB API provides a minimal standard for working with databases using Python structures and syntax wherever possible. This API includes the following:

* Importing the API module.
* Acquiring a connection with the database.
* Issuing SQL statements and stored procedures.
* Closing the connection

MySQLdb is an interface for connecting to a MySQL database server from Python. It implements the Python Database API v2.0 and is built on top of the MySQL C API.

The Natural Language toolkit, or more commonly NLTK, is a suite of libraries and programs for symbolic and statistical natural language processing (NLP) for the Python programming language. It segments the sentences and tags these according to the part of speech. The program can be designed in the language that is suitable to both the tablet device and the Personal Computers and which properly incorporates the Natural Language Processing Toolkit.

Python3 is installed from Python.org and by using pip installer we install the required wheel files for MySQLdb, flask and nltk.

DATA WAREHOUSE

Enterprise Data warehouse contains all information about subjects. The Data Mart is the departmental subset that focuses on selected subjects. The aforementioned artificial intelligence mechanism forms the Decision Support System (DSS). We perform drill downs and then maintain meta-data for our warehouse. The data about the fee structure, marks and salary forms the operational data. The user guidance support forms the informational data. We make use of the Structured Query Language to add the primitive data and the feedback which is in short the operational data. The informational data is added using Python snippets as mentioned above. We use MySQL Database Management system as our Relational Database Management System for the warehouse.

DATA MINING

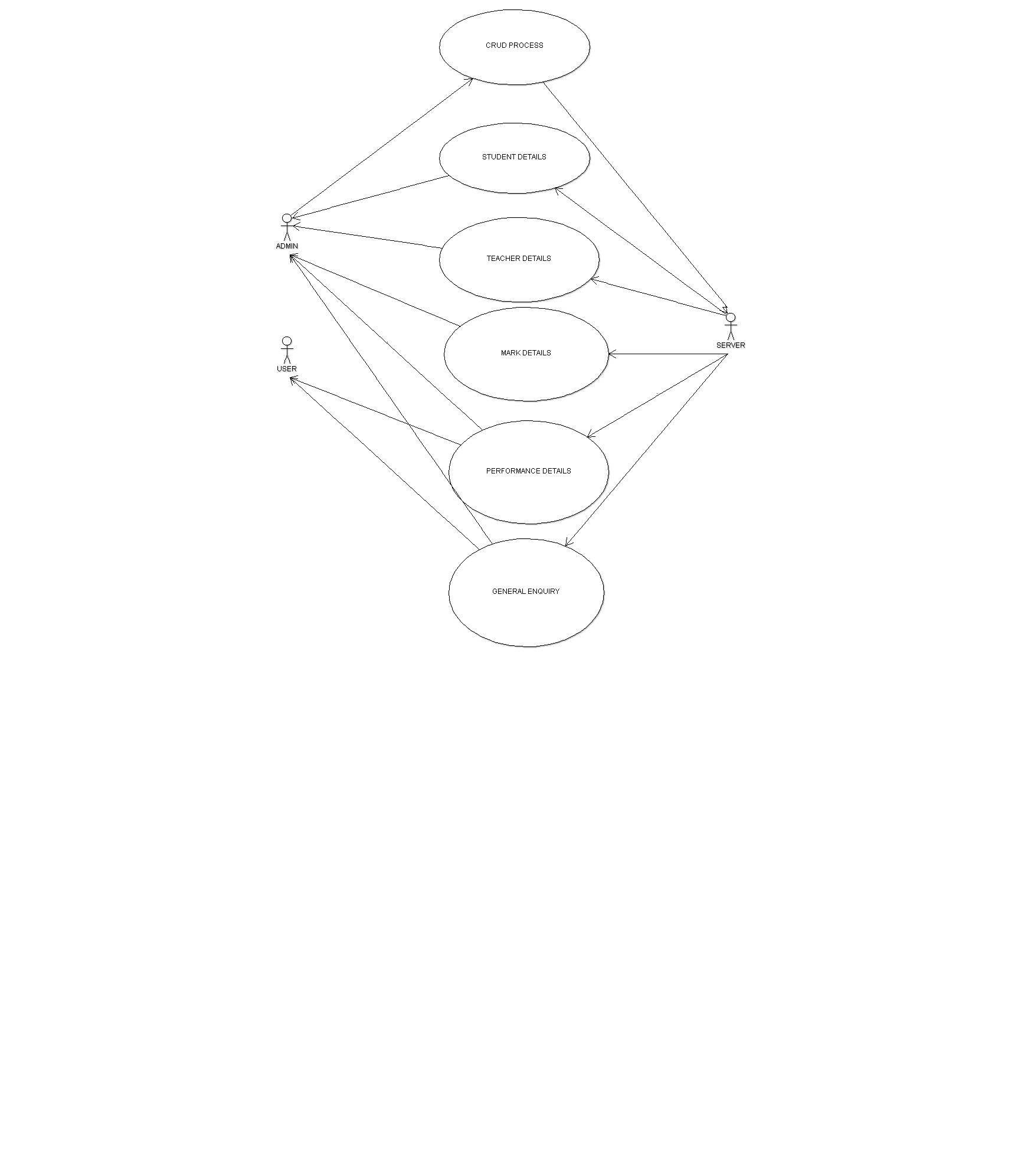
Data mining refers to extracting or mining knowledge from large amounts of data. The tasks involved in Data Mining: characterization, discrimination, association, classification, prediction and clustering analysis can be implemented in the Python programming Language. Main objective is the identification of a particular pattern to aid interpretation. This pattern may be either subjective or objective. We include functionalities to perform preprocessing of data to present the final data, the decision tree and also the probability prediction using Bayesian network which is a representation of variables or data and their conditional dependencies. Outliers can be implemented for safety of privacy [2].

SYSTEM DESIGN

Systems design is the process or art of defining the architecture, components, hardware & software modules, interfaces, and data for a computer system to satisfy specified requirements. This could be seen as the application of system theory to computing and product development. Some overlap with the discipline of systems analysis appears inevitable. Designs tools such as Unified Modeling Language, UML now address some of the issues of computer systems design and interfacing.

SYSTEM ARCHITECTURE

System architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system.



IMPLEMENTATION:

Implementation is the stage in the project where the theoretical design is turned into a working system. The most critical stage is achieving a successful system and in increasing the reliability of the new system. It involves careful planning, investigation of the current system and its constraints on implementation, design of methods to achieve the changeover and an evaluation of change over methods.

The following working model is explained keeping in mind its implementation using the Python programming language.

MODULE DESCRIPTION

DATABASE CONSTRUCTION AND DEPLOYMENT (MySQLdb module)

To work with the data, the priory is to first upload the raw data in the server. What better way than MySQL server to achieve the goal. But sometimes manually entering large number of queries is difficult and time consuming. So for that we provide a module which helps in easy insert, update and delete with the records. This module is a Python code having a set of predefined SQL statements executed based upon the user choices after appropriately establishing a connection. Plus the exceptions are neatly handled in case user enters an inappropriate data.

CONSTRUCTION OF DATA MINING MODULE THROUGH NLP ALGORITHM (nltk module)

In this module we go about in constructing a proper and effective NLP module which does the syntactic, semantic and pragmatic analysis to mine the required data we need. The nltk module in Python provides function to tokenize a whole sentence into a collection of words. And then it aids in Data mining by carrying out a repeated lookup in the predefined lists which acts as the dictionary here. After the required keywords are shortlisted, the appropriate action is executed entirely through the Python code.

SERVER DEPLOYMENT MODULE (flask module)

We have obtained the required mined data. But it is important to make the information available to all the users. And since we want all users to have access to it, we deploy it to the server at Localhost using Flask library in Python. Python scripts are fast and flask being a light weight framework supports the speed of execution at run time. A flask program creates a module for the application, a set of definitions for the URL mappings of the application and finally a main class which encloses a function to start the application at local host in the server.

IMPLEMENTATION DETAILS

The input is given either in the form of text as an administrator to upload the initial data and voice for other normal users. For the administrator module, the person assigned as admin is required only to input the required data and not the strenuous commands with it as Python provides an easily understandable user friendly interface which can be used by people with basic knowledge in using computers. Once the required data is inserted, deleted or updated, it is ready for mining.

The user or the admin fires a request for mining the desired data. He provides the input as voice recorded or to put in proper words recognized with the help of the Web Speech API in HTML 5. The obtained text is passed as an argument to the URL containing the code for mining the data. This code effectively follows the NLP algorithm. The syntactic analysis is carried out to tokenize the sentences we speak. Then the semantic analysis, groups the necessary tokens to derive the real meaning it intends to put across. Finally the results of voice goes to the Pragmatic Process module which decides what has to be done for the user request. In this case, the predefined lists will check out if such user request can be handled.

Then the control moves to the Data Warehouse containing a large number of databases or Data Marts. The tuples which we require is fetched from appropriate Data Marts. This raw data or tuples has to be preprocessed before analyzing the pattern between different relevant data. In this we perform cleaning, integration, transformation, reduction and finally discretion. The database manipulation and analysis is handled by the MySQL Relational Database Management System.

Now we utilize the database API of Python to connect the MySQL with the Python code. Now we are utilizing Python codes instead of Procedural and trigger statements of MySQL as Python includes easier and more functionalities. The Mining Process takes with the queried data as input and it involves several internal process to obtain the required pattern. Data characterization involves summarization of the general characteristics or features of target data. In Data Discrimination, we perform comparison of the general features of target data with the general features of data from one or a set of contrasting classes. Then we establish a proper association relationship between different relevant datum considered. And then we classify the data based on numerical and probabilistic computations. Finally we cluster the data to include only those which we need to draw up a pattern.

With the help of that pattern (set of function calls)we attempt to trace a decision tree which is a representation of the control flow across different functions to present the needed result.

The result obtained has to rendered in a proper format for which we utilize Bootstrap which is a set of predefined CSS codes for ensuring a neat and responsive design. The output is presented as an AISN response and all the outputs related to the input are presented without any omission.

NLP ALGORITHM

The new algorithm introduced for effective data mining from voice input is the Natural Language Processing Algorithm. There are three major aspects of any natural language understanding theory: The syntax describes the form of the language. It is usually specified by a grammar. The semantics provides the meaning of the utterances or sentences of the language. The pragmatic component explains how the utterances relate to the world. To understand language, an agent should consider more than the sentence; it has to take into account the context of the sentence, the state of the world, the goals of the speaker and the listener, special conventions, and the like.

DIAGRAM --

The efficiency of Natural Language Processing in real world implementation is not exactly 100 percent perfect. But it is just about accurate for the system to understand and generate an algorithm. The easily or mostly solvable problems include Spam Detection, tagging, named entity recognition. Further problems include Sentiment analysis, Co-reference resolution problem, and Word sense disambiguation problem, Parsing, Machine Translation and Information Translation. These problems are solvable if proper research is done and corresponding implementation is included. Some of the problems which are almost impossible to solve are summarization of input and implementation of a dialog system that prompts a related query to the input if the input is ambiguous. Our system simply reacts by saying something like “I do not understand you. Come again”.

RESULTS AND DISCUSSION

RESULTS

Suppose a department Head wants to set up the time-table for the new semester, here is how the process takes place in the application. The input would be "Please schedule the time-table for CSE". Due to the lack of information at the semantic phase, all available framed timetables will be displayed. From the tokenization, the key words identified were "time-table", "CSE". Now referring to the dictionary, the appropriate case was located by the agent and the input was understandable. Now it fetches the mark table and the passstatus table from the Data Warehouse. After prior preprocessing, we extract the tuples needed for that particular branch and year from the first table and then we extract the previous year results of Computer Science and Engineering alone from the second table. In the mining process we obtain a pattern relating the difficulty of the subjects to the results and subsequently the allocation of number of classes. If some subjects have almost the same level of difficulty, then probabilistic calculations are made which decides which subject should be given the most number of classes after the comparison with other subjects. The data virtual tree is pruned to contain only the subject information and the related pattern. Finally a subject can be allocated at different times based on the conditional probability. Here is a decision tree snippet to illustrate class allocation.

Assigning difficulty,

IF (result > 80) THEN difficulty = "easy";

ELSE IF (60 < result <=80) THEN difficulty = "medium";

ELSE difficulty= "hard";

Same level of Difficulty issue,

IF difficulty= "hard" and probability="yes" THEN

num-class=5;

ELSE IF difficulty= "hard" and probability="no"

THEN num-class=4;

ELSE IF difficulty= "medium" and probability="yes"

THEN num-class=4;

ELSE IF difficulty= "medium" and probability="no"

THEN num-class=3;

ELSE num-class=3;

Another instance would be the salary increment calculator of the faculties based on their average and pass rate. Consider the marks updated periodically, then finally she will have a particular rating at the end of the semester evaluated based on the pass percentage and average. The respective fields are fetched from the Warehouse and the average of the above 2 is calculated during the classification phase of the mining process. Then the following Decision Tree Snippet shows what to do:

average calculation,

current\_average= ( average\_in\_subject + pass\_percentage )/2 ;

salary increment calculation,

increment=( current\_average-previous\_average );

IF increment>0 THEN salary+=increment\*100;

ELSE salary+= 0;

We can easily set a easy or a moderate or a tough question paper by utilizing the Bayesian Network mechanism in the Prediction phase. It is based on Conditional probability relating the probability of maximum or minimum occurrence of question with the probability of occurrence of other similar questions of the same discipline.

SCREENSHOTS

CONCLUSION

This application is generic and can be extended to different areas such as food chains, tourism, defense, agriculture and the like. Future likely addition to the above application would be the sensors. For instance to automatically sense and record the attendance saving further time. This would lead to the creation of a powerful Internet of Things tool which is the need of the hour at the moment.

FUTURE WORK

Some of the listed plans for future are as below:

1. Implementing a machine learning module again using Python utilising PyBrain library.

2. Providing a hybrid application portable across multiple mobile devices using Apache Cordova.

3. Making insertion,deletion and updation in admin side completely artificial Intelligent.

REFERENCES

[1] Steven Bird, Ewan Klein and Edward Loper, “Natural Language Processing with Python,” O'REILLY.

[2] Alex Berson and Stephen J. Smith, “ Data Warehousing, Data Mining & OLAP”, Tata McGraw – Hill Edition, Tenth Reprint, 2007.

[3] Jiawei Han and Micheline Kamber, “Data Mining Concepts and Techniques”, Elsevier, 2007.

[4] Y. C. Huei, "Benefits and introduction to python programming for freshmore students using inexpensive robots," Teaching, Assessment and Learning (TALE), 2014 International Conference on, Wellington, 2014, pp. 12-17. doi: 10.1109/TALE.2014.7062611

[5] V. Aswini and S. K. Lavanya, "Pattern discovery for text mining," Computation of Power, Energy, Information and Communication (ICCPEIC), 2014 International Conference on, Chennai, 2014, pp. 412-416. doi: 10.1109/ICCPEIC.2014.6915399