Title – Descriptive Online Examination Automation System Using Natural Language Processing

Guide - Mr. Madar Bandu (Assistant Professor)

Department of Computer Science and Engineering, Anurag Group of Institutions

Madarbanducse@cvsr.ac.in

M Suhas Rao

Department of Computer Science and Engineering, Anurag Group of Institutions

18h61a05m2@cvsr.ac.in

Maddhi Shashidhar Reddy

Department of Computer Science and Engineering, Anurag Group of Institutions

18h61a05m3@cvsr.ac.in

Mangala Balakrishna

Department of Computer Science and Engineering, Anurag Group of Institutions

18h61a05m5@cvsr.ac.in

ABSTRACT

This world has a large number of test sites that can be distributed over several servers used for online trials for a variety of purposes, among which a few may include writing entrance exams, or Olympics at the national and international level and even a few portals are designed to serve setting goals, but we have seen that, in general, all portals are designed for the same behaviour that contains a few choice questions. Here our goal is not to build the technology that is currently available, but rather a very different project that solves the big problem of online testing. Here we use a descriptive online test program. Many selection questions are easy to deal with as it has a question, other options and discipline within the equal question that keeps the right choice within the website. While in the case of descriptive questions, this is not always the case. Provides the principles of Natural Language Processing to provide feedback marks. Solutions are nothing but strings, and the model's job is to do a little work on the solution series to provide accurate marks on the solutions written using the examiner. Facts are collected in an online descriptive test machine. In addition, its revised and designed version assigns accurate marks to the solution to the question.

Keywords: Machine learning, Natural language processing, Semantic Matching.

INTRODUCTION

Motivation

Despite the major disadvantages of the offline exam system, the traditional exam system will not be replaced, as the new online exam system only has multiple choice type questions, although most exams have detailed questions, which have multiple choice answers. They do not work and, therefore, it is not convenient and effective to replace them on a large scale. We all know that if we want to get rid of the universally accepted system, the new system must not only be better, but also change in quality so that companies can accept it.

Problem Description

In the proposed model, we are taking the online examination system to a new level by enabling the examinee to write descriptive answers which will be self-evaluated, i.e., automating the entire offline examination system with computing efficiency without any human error. This can be done using Text analysis and Semantic matching using NLP techniques. The assessed answers will be stored in the database and can be accessed anytime and a special student profile will be maintained for better evaluation of the student.

Purpose of the Project

This is a big boost to the online examination system as it allows it to overcome its biggest disadvantage and helps the online examination system to raise its paw in the mid-year or annual examinations conducted for school or college evaluations. This system frees teachers and professors from the burden of checking copies and makes them more productive with their time in teaching, which also eliminates bias in the review of answer scripts and takes up leasing space. The copies will not be scanned due to any human error and full marks will be allotted as per the written procedure by the examiner and there will be more or less room for acquisition, which will help in resource management as it is a frequent cut. It has higher efficiency over time as it gives instant results and is safer and more reliable.

EXPERMENTAL

Python

Python is one of the most powerful languages currently in the world in terms of packages and libraries that provide data analysis and machine learning process, in this project we need many libraries such as numpy, pandas, nltk.

PyCharm

PyCharm is an IDE used for computer programs, in particular the Python language.

Natural Language Tool Kit

The purpose of the Descriptive Online Examination Automation System Using Natural Language Processing is automating the process of examination by analysing the questions and answers and calculating the marks. Data pre-processing, the initial part of the project is to understand implementation and usage of various python inbuilt modules. The above process helps us to understand why different modules are helpful rather than having to implement those functions from the beginning by the developer. These various modules provide better code representation and user understanding.

Exploratory data analysis, first step in this to apply a Text analysis algorithm i.e., BERT and TF-IDF algorithms which for evaluating answers for questions that are descriptive in nature. Keywords are stored for every question and on the basis their occurrence in the answer string, the examinee is allotted marks. In order to perform this in Python language, there exists a library that makes it a little easier for the algorithm developers to perform string manipulations. The name of the library used is NLTK which is specifically designed for python to work on NLP. It is discussed in detail in the section where the algorithm is discussed.

Talking about the technology used in order to build such a model for evaluating descriptive answers, NLP or Natural Language Processing is has a great role to play. NLP can do a lot of innovative jobs like predicting if a message or an email is a spam or a ham, the quality search that we can do on shopping websites like amazon and flipkart in order to search for different categories of items that include kitchen utensils, electronics gadget, apparels, food items and much more such products that are available online. The basic idea was that did anyone ever think of knowing how these search bars or how these ham spam classifications work? The answer to this question is that rarest of the rare people have tried getting into this and tried to know what the mechanism or the back-end work in order to give such powerful search results and such predictive classification techniques. For those who are not aware of the mechanism behind this, it's all just about playing with strings of characters, numbers and special characters or what we call as string manipulations to arrive at such results.

Existing System

This world has seen a lot of exam portals which are deployed on multiple servers which are used to conduct online exams for various purposes for Olympiads and some are i.e., designed to conduct an examination for placement purposes. But what we have observed is that most of the portals are designed to conduct tests that consist of multiple-choice questions.

Disadvantages

Offline examination system is low efficient because of the resources been used and the time consumed and the existing online examination system features only multiple-choice type of questions.

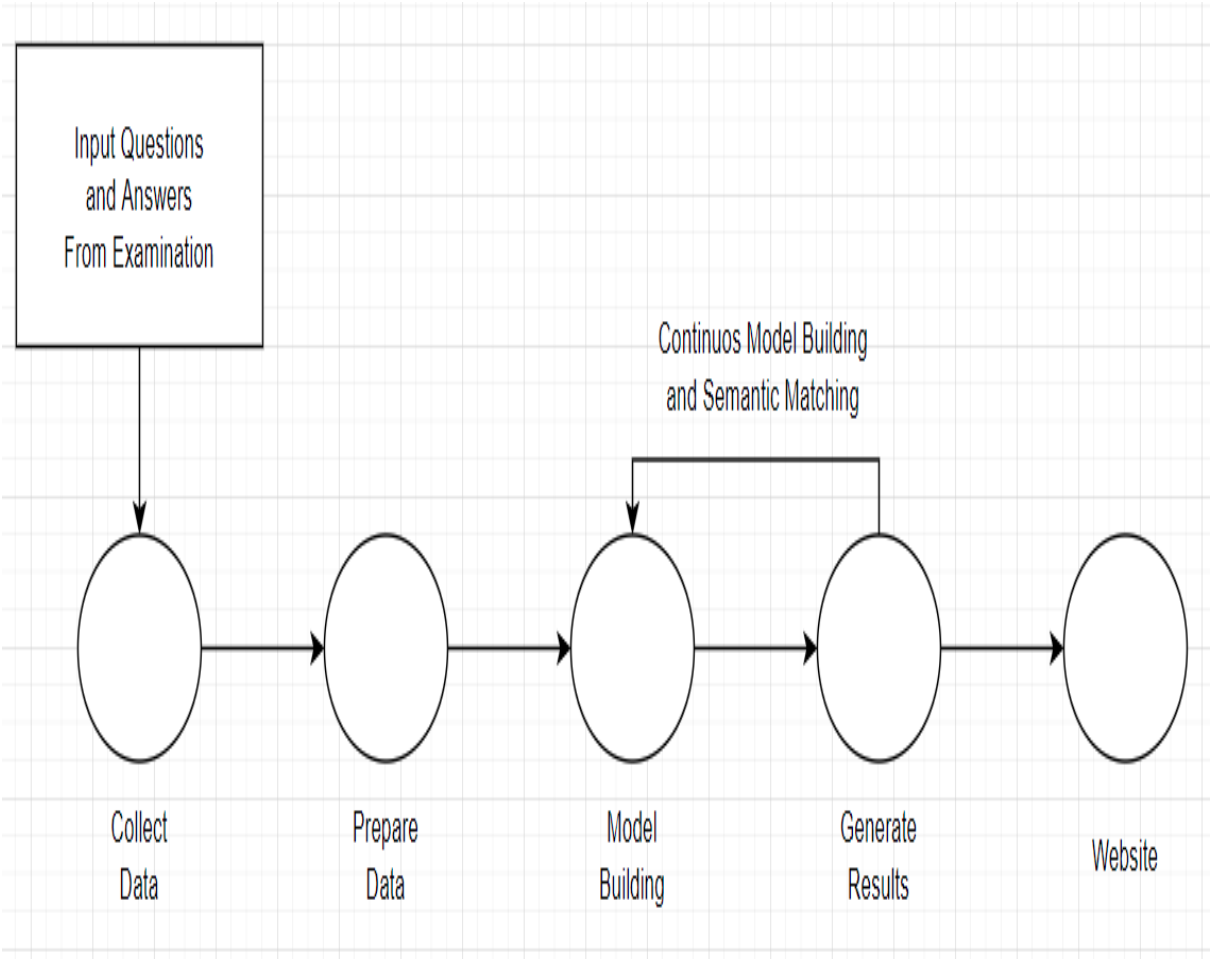
Proposed System

The proposed model takes the online examination system to a new level by enabling the examinee to write descriptive answers which will be automatically evaluated. The evaluated answers will be stored in the database and can be accessed anytime and a special student profile will be maintained for better evaluation of the student. Talking about the technique used for building such models for evaluating descriptive answers, NLP or Natural Language Processing has a big role to play.

Online examination system is highly efficient and answers are evaluated at that itself and student can see the solutions and can correct the mistakes or errors committed while appearing for the exams.

RESULTS AND DISCUSSION

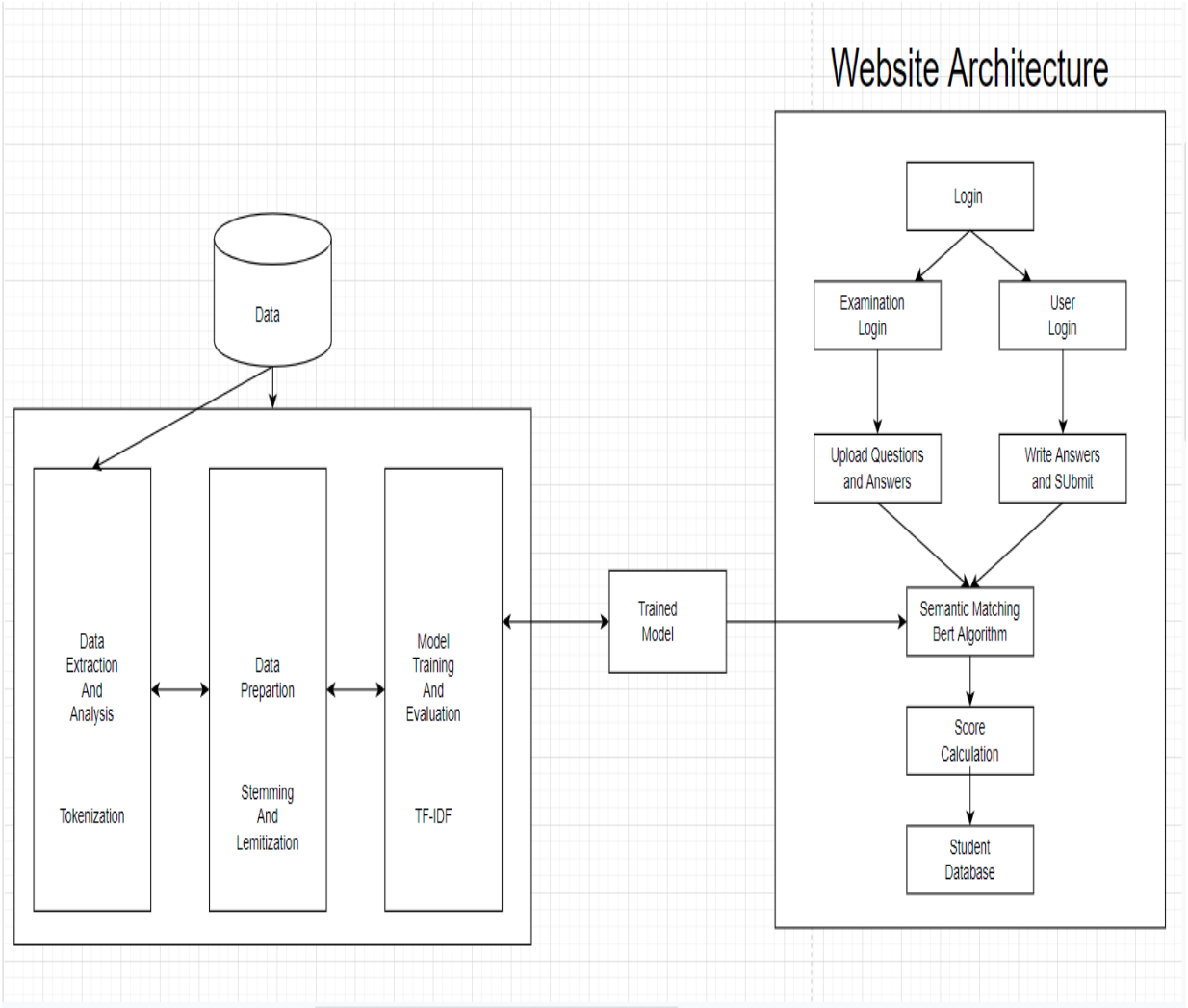
The method of Descriptive Online Examination Automation is divided into different stages, the process includes gathering relevant data, prepare data, model building and Generate results and upload it to the website to obtain the results.



Collecting data is the first step of the process flow which includes defining project, setting up the machine environment suitable for the development requirements and later understanding the data using different python libraries and machine learning techniques. Data Cleaning need to done on the data collected so that the analysis be very accurate for perfect results.

The questions and answers are upload by the examination which will be collected and the data extraction and analysis is done on the data using the tokenization algorithm

After the data extraction and cleaning is done we will be doing Data preparation using Stemming and Lemmitization algorithm where the tokens generated are separated and given separate branches based on the category of the words, then the model training is done by using algorithms TF-IDF to generate the specific keywords and after using the BERT algorithm the semantic matching is done.



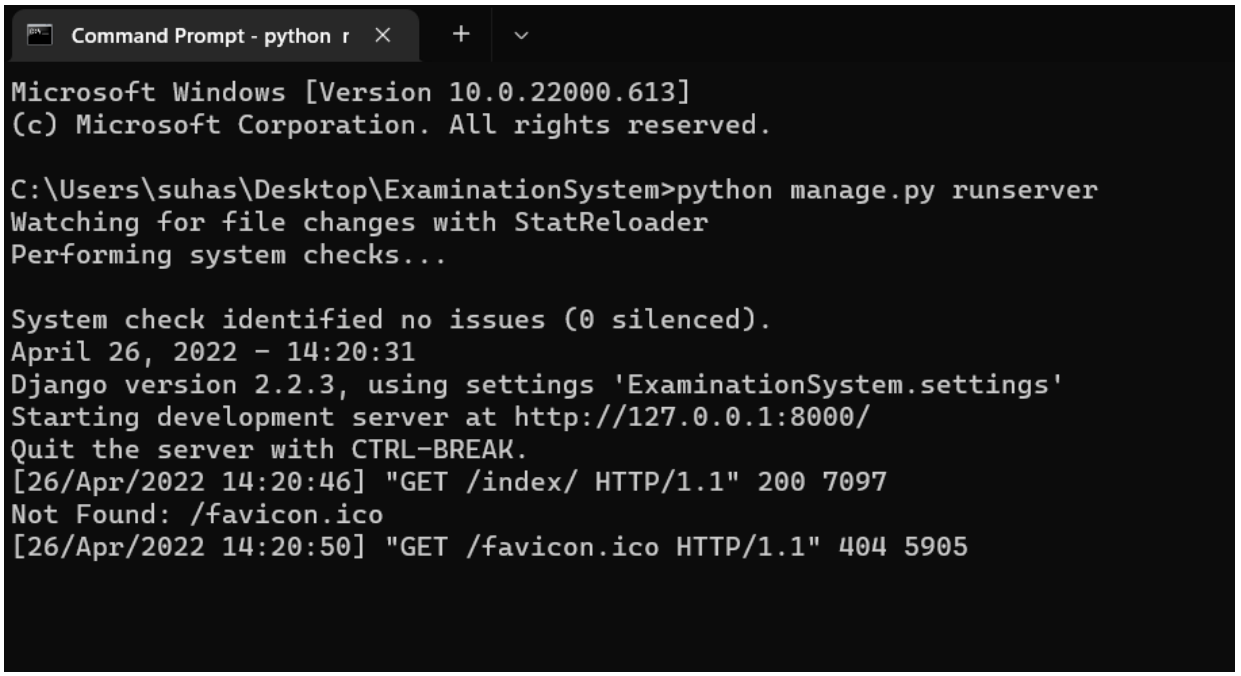
EXECUTION PROCESS

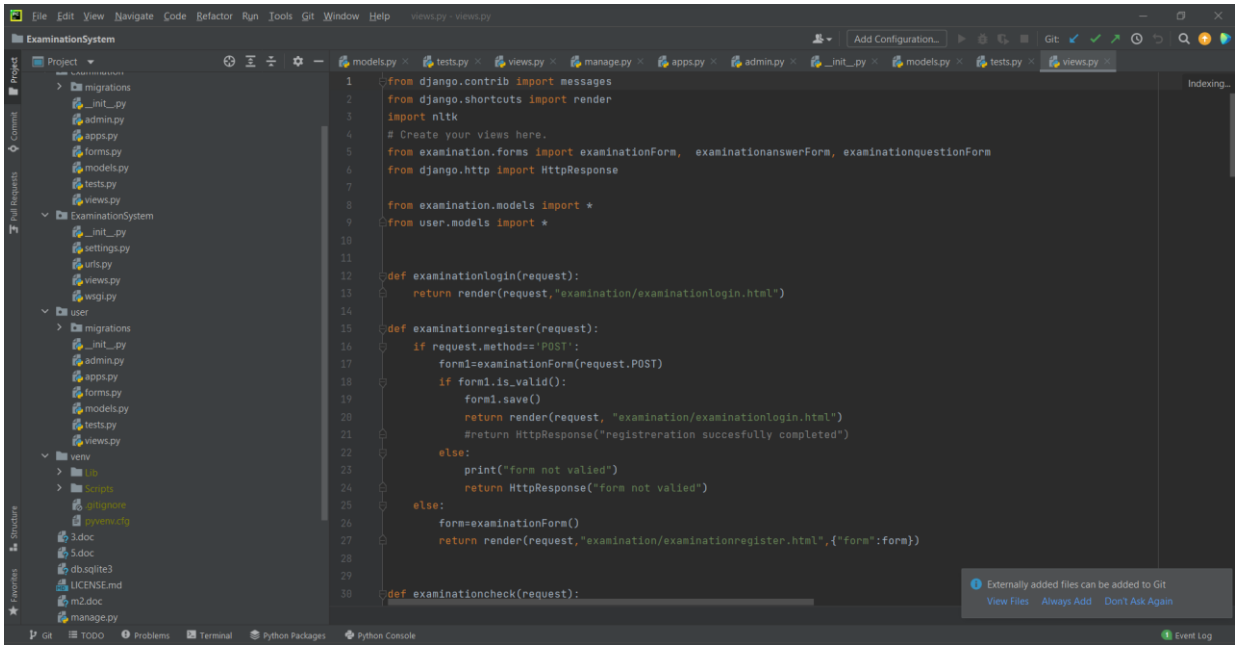
The different execution process includes different steps like preprocessing of the data, data exploration and Text pattern analysis and semantic matching and finally showing the output in the app window using the website.

RUNNING THE MODEL

pip install -r requirements.txt
python
>>import nltk
>>nltk.download()

- python manage.py runserver
- copy the url generated and paste it in browser
- ip-address/index in browser





READING THE INPUT

As soon as you click on command python manage.py runserver in terminal the system gets loaded and the output window will be appeared now you can upload the text file and do the analyses process.

OUTPUT

The Below Figure Shows the output of the execution process

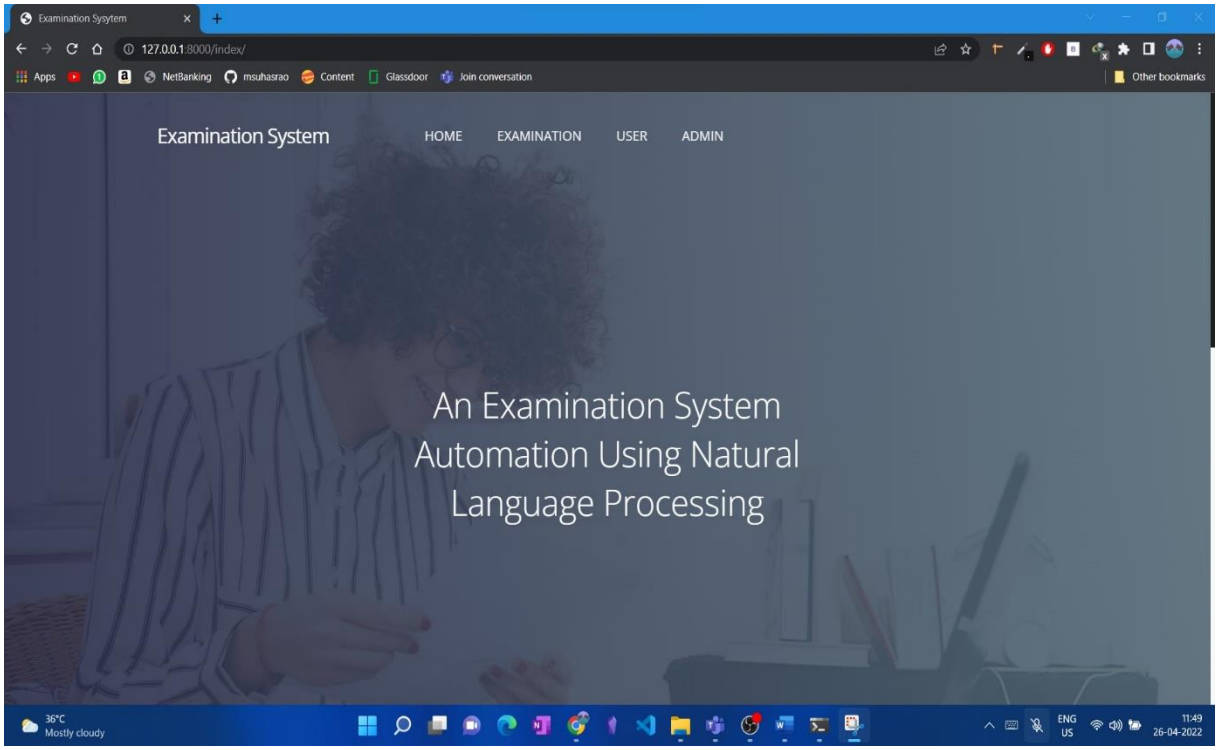
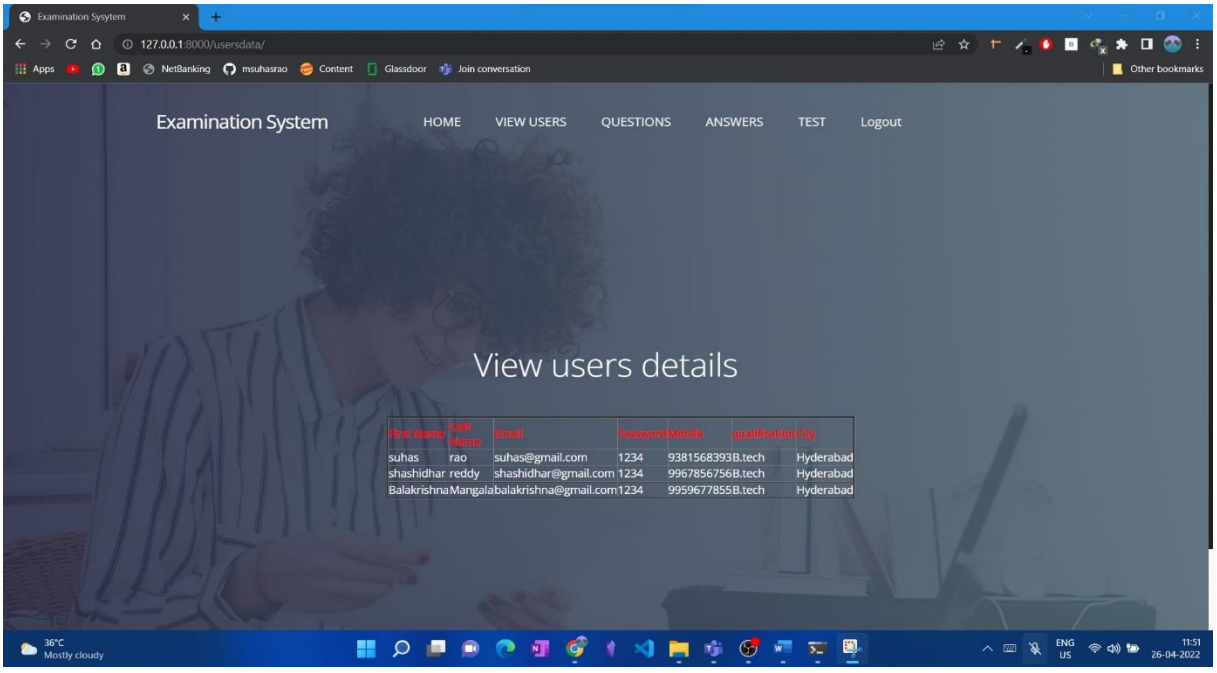
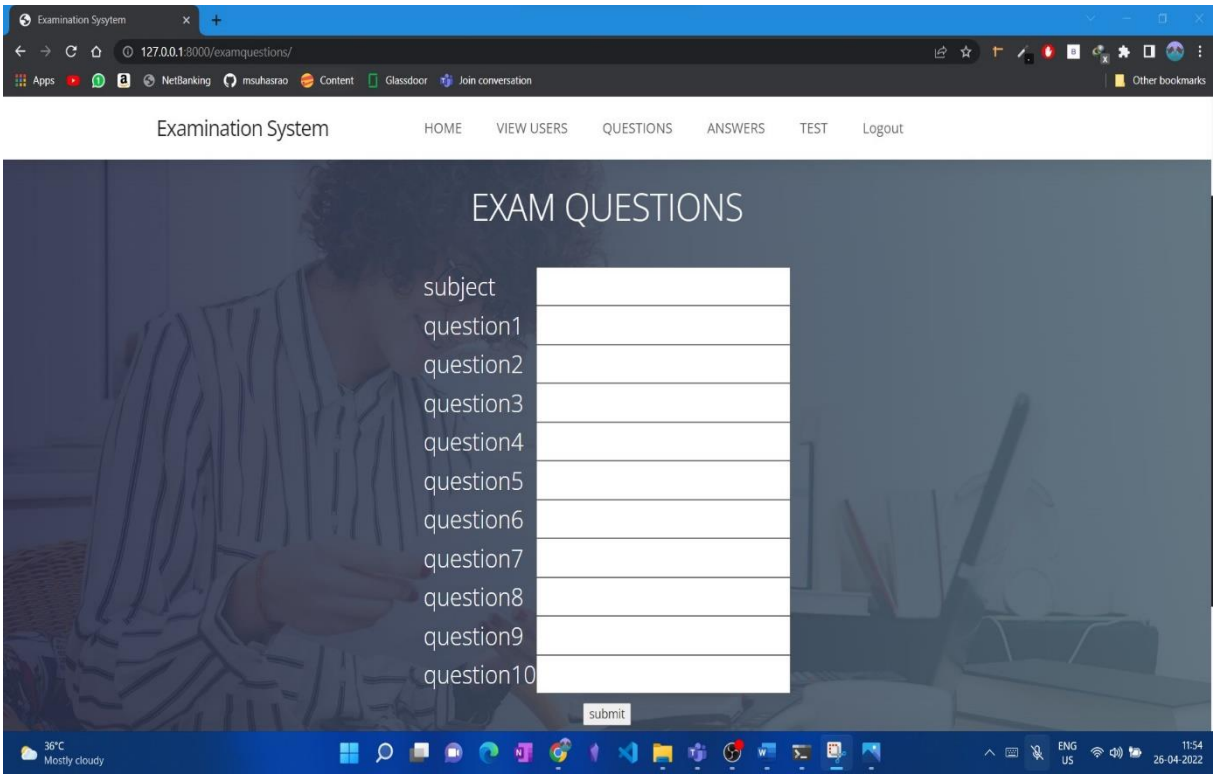


Fig Home page

It consists of 3 tabs examination, user and admin





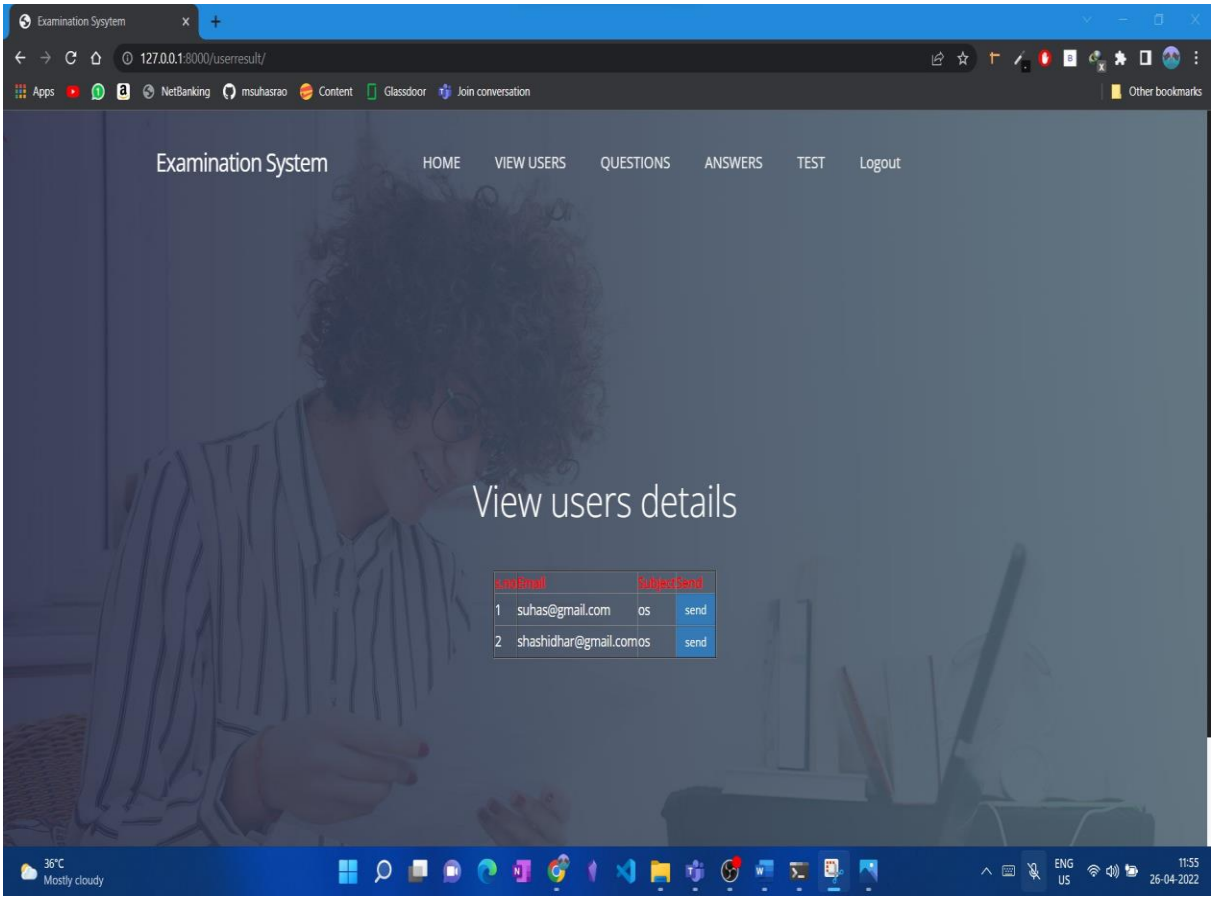


Fig Test result page

CONCLUSION

In conclusion, It can be seen by conducting tests using such an algorithm at regular intervals that one can determine the trend in the marks obtained by different students and we can give them an analyzed report on the different subjects they need to focus on for which they are weak. With the existing data, we can also implement a predictive machine learning model on the data so that it can predict marks that the students will score in the future. It is observed that students mainly study those subjects that are placement oriented or which are required for placement purpose only. While students neglect the subjects of their core domain. Deep knowledge in the domain is required as it is of no use to study if you do not have a core domain knowledge. So it can help students get quality knowledge as everything will be digital and there will be no cumbersome process of conducting a pen-paper test. Also, answers are evaluated at that moment itself and the student can see the solutions and can correct the mistakes or errors committed while appearing for the exam. databases of social media and online social media platforms of various sources that host different companies. Online Examination System is widely used as compared to other exams. These bugs must be identified and solved for improving quality of software. So, in future we can develop more secure software by using advanced technologies, by introducing new security systems using biometrics, we can identify the student's true identity by analyzing digital signature or by finger print or by capturing images and also by using online live proctoring.

REFERENCE

[1] K. Jayakodi, M. Bhandara and I. Perera "An automatic classifier for exam questions in Engineering: A process for Bloom's taxonomy", IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), (2015)

[2] N. Ishikawa, K. Umemoto, Y. Watanabe, Y. Okada, R. Nishimura and M. Murata "Detection of users suspected of using multiple user accounts and manipulating evaluations in a community site", IEEE Proceedings of the 6th International Conference on Natural Language Processing and Knowledge

[3] B. Kaur, and S. Jain "Keyword extraction using machine learning approaches", IEEE 3rd International Conference on Advances in Computing, Communication & Automation (ICACCA) (Fall), (2017)

[4] R. P. Futrelle, J. Satterley, and T. McCormack “NLP-NG — A new NLP system for biomedical text analysis”, IEEE International Conference on Bioinformatics and Biomedicine Workshop, (2009)

[5] M. Revathy, and M. L. Madhavu” Efficient author community generation on Nlp based relevance feature detection”, IEEE International Conference on Circuit, Power and Computing Technologies (ICCPCT), (2017)

[6] W. Nei, Y. Wu, D. Hu, L. Wang, and Y. Li” Data Management and Analysis of Intelligent Examination Scoring System of Simulation Training System”, IEEE 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, (2013)

[7] G. Zhang, and H. Ke” Design of Paperless Examination System for Principles of Database Systems”, IEEE International Conference on Research Challenges in Computer Science, (2009)