AN EXAMINATION SYSTEM AUTOMATION USING   NATURAL LANGUAGE PROCESSING

A.   Amrutavalli[1], A.Swathi [2], P.Rishitha Jahnavi [3], J.Raveendra [4], T.Venkata Subbaiah [5]

[1] *Asst. Professor, Krishna Chaitanya Institute of Technology & Sciences , Markapur, A.P, India*
[2,3,4,5] *Scholar, Krishna Chaitanya Institute of Technology &Sciences , Markapur, India*

ABSTRACT:

This world has seen a lot many examination portals that are deployed over several servers which are used to conduct online examination for various purposes among which some may include conducting a test for entrance examinations, or Olympiads at national and international level and while some portals are designed to conduct a test for placement purposes.But what we have seen is that mostly all the portals are designed to conduct tests that contain multiple choice questions. Here our aim is not to work on the technology that is already existing,rather some technology that is very rare. Here we talk of the descriptive online examination system. Multiple choice questions are easy to deal as they have a question, a few options and a field in the same question that stores the correct option in the database. While in the case of descriptive questions it is not so. It brings in or uses the concepts of Natural Language Processing or NLP to assign marks to answers. Answers are nothing but strings and the job ofthe model is to do some operations on the answer string such that it can assign the correct marks to answers written by the examinee. The data is basically collected from a descriptive online examination system. Further, it is analyzed and the designed model assigns accurate marks to the answers for the question. The back-end is written in Python where the web framework usedis Django, the library used for Natural Language Processing includes NLTK and for database purpose, SQLite version 3 is used, while for the front-end HTML version-5, CSS version-3, Bootstrap and JavaScript is used.

Keywords—Exam System; SQLite3; Django; Descriptive System; Natural Language Processing, Python; NLTK

[1] INTRODUCTION

We come to hear news from around the globe that a particular exam was conducted for a job or for a college or examination in schools and the result was published after some time, while this is a good way to conduct an exam but it is inefficient with respect to the current worldwhere automation is the future. The examination system relies on manual work from printing to transporting the paper to the examination hall, then invigilation and the most tedious task of checking the answer sheets which is a huge mess for any examiner which sometimes leads to resource loss. Also, we hear news about paper leaks and answer sheet being lost in the

transporting process. The manual checking process will always have that human error based on certain factors like biasing, the mood of the examiner, target completion and much more such factors. Also if we take account of all the paper wastage and the stationary waste which harms our environment leading to do more bad then good as the enormous amount of trees being chopped off across the world for the process. This helps us understand that the offline examination system is not cost effective or time efficient, resources are also wasted in the process and moreover we all know that resources are scarce in nature and we need to utilize it efficiently to get the maximum output of it.

While the offline examination system has a big disadvantage but are not getting replaced at a bigger scale because new online examination system features only multiple choice type of question's while most of the exams contain descriptive question for which multiple choice answers do not work and hence they are not that compatible and efficient to replace it at a larger level. We all know that if we have to remove a universally accepted system, the new system should not be just good, rather it should be able to make a quality difference so that the organizations accept it. While there are some examination system and they are good at evaluating the answers but they have little to no scope for the descriptive ones and the analysisis not well implemented to get meaningful results. Even most famous of them just have a simple system of storing the correct options in the database and just matching the correct option with it to calculate the result.

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 get evaluated on their own i.e automating the entire offline examination system with the efficiency of computing having no human error involved, this can be done using NLP or Natural Language Processing. Theevaluated answers will be stored in the database and they can be viewed anytime and a particularstudent profile will be maintained for better evaluation of the student.

This will be a huge boost to the online examination system as this will allow it to overcome its biggest con and it will also help the online examination system to stretch its paw even in the half-yearly or annual examination conducted by schools or college for evaluating theprofile of the student. This will have instant benefits like the system will relieve the burden of the teachers and professors of checking copies and in return they can be more productive with their time in teaching things, this will also eliminate biasing in answer script checking and will have leased space for any human error as copies would not be scanned and the entire marks will be allotted according to the way answers are written by the examinee while he was on and there will be little to no space for acquisition, it will help in resource management as this will cut corners on stationery products, it will also have greater efficiency with respect to time as it will produce instantaneous results and will be more secure and reliable.

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 cando 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 www.amazon.in and www.flipkart.com 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 hamspam 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 notaware 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.

For example when a product is stored in the database of the online shopping website, what happens is that additional keywords are stored for them such that they can be searched using those keywords. The string input by the customer in the search bar contains keywords using which products are searched and displayed on the customer dashboard. Even the back-end for such applications contains quality search algorithms that are out of the scope of this paper's discussion which is on the model of the descriptive online examination system. Similarly, the scenario is quite the same here 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.

[2] LITERATURE SURVEY

Assessment is an essential activity to achieve the objective of the course being taught and to improve the teaching and learning process. There are several educational taxonomies that can be used to assess the efficacy of assessment in engineering learning by aligning the assessment tasks in line with the intended learning outcomes and teaching and learning activities. This research is focused on using a learning taxonomy that fits well for computer science and engineering to categorize and assign weights to exam questions according to the taxonomy levels. Existing Natural Language Processing (NLP) techniques, Word net similarity algorithms with NLTK and Word net package were used and a new set of rules were developed to identify the category and the weight for each exam question according to Bloom's taxonomy. Using the result the evaluators can analyze and design the question papers to measure the student knowledge from various aspects and levels. Prior evaluation was conducted to identify most suitable NLP preprocessing techniques to the context. A sample set of end semester examination questions of the Department of Computer science and Engineering (CSE), University of Moratuwa was used to evaluate the accuracy of the question classification; weight assignment and the main category assignment were validated against the manual classification by a domain expert. The outcome of classification is a set of weights assigned under each taxonomy category, indicating the likelihood of a question to fall into a certain category. The highest weight category was considered as the main category of the exam question. Accordingto the generated rule set the accuracy of detecting the correct main category of a question is 82%.

Some users in a community site abuse the anonymity and attempt to manipulate communications in a community site. These users and their submissions discourage other users, keep them from retrieving good communication records, and decrease the credibility of the communication site. To solve this problem, we conducted an experimental study to detect users suspected of using multiple user accounts and manipulating evaluations in a community site. In this study, we used messages in the data of Yahoo! chiebukuro for data training andexamination. This study found journalists use government sites most often to retrieve information. Problems include difficulty with verification, unreliable information and lack of contact information. Paperless examination is an important role of modern education, which can effectively reduce the teachers' workload and improve work efficiency. However, the current paperless examination system mainly deals with the objective questions, it is almost impossible to deal with subjective questions such as programming languages, particular in SQL. There is no such practical system as far as know. This article describes a novel SQL-based paperless examination system, including objective questions as well as SQL programming questions.

Computer greatly influences our educational environment. Over the last years, automatic computer examination systems have been widely used for computer-based tests. But these systems are based on traditional question-answer examination style which is not fit for the sequenced test. The sequenced test should consider the context of the examinee, e.g. the order ofquestions or the permissions of the examinee, to grade an examiner. In this paper, we propose an effective and practical automatic examination architecture based on task. The task is abstracted from the examination process and can meet the requests of the sequenced test, such as order and dependency. At the end of the paper, we implement an automatic examination system based on task for the stake test which proves quite efficient in practice.

In Existing System, This world has seen a lot many examination portals that are deployed over several servers which are used to conduct online examination for various purposes among which some may include conducting a test for entrance examinations, or Olympiads at national and international level and while some portals are designed to conduct a test for placement purposes.But what we have seen is that mostly all the portals are designed to conduct tests that contain multiple choice questions. In Existing System have some Disadvantages are offline examination system, online examination system features only multiple choice type of question's, and Low Efficiency

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 get evaluated on their own The evaluated answers will be stored in the database and they can be viewed anytime and a particularstudent profile will be maintained for better

evaluation of the student. 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. In the proposed model advantages are High Efficiency, online descriptive examinations system and Answers are evaluated at that moment itself and the student can see the solutions and cancorrect the mistakes or errors committed while appearing for the exam.
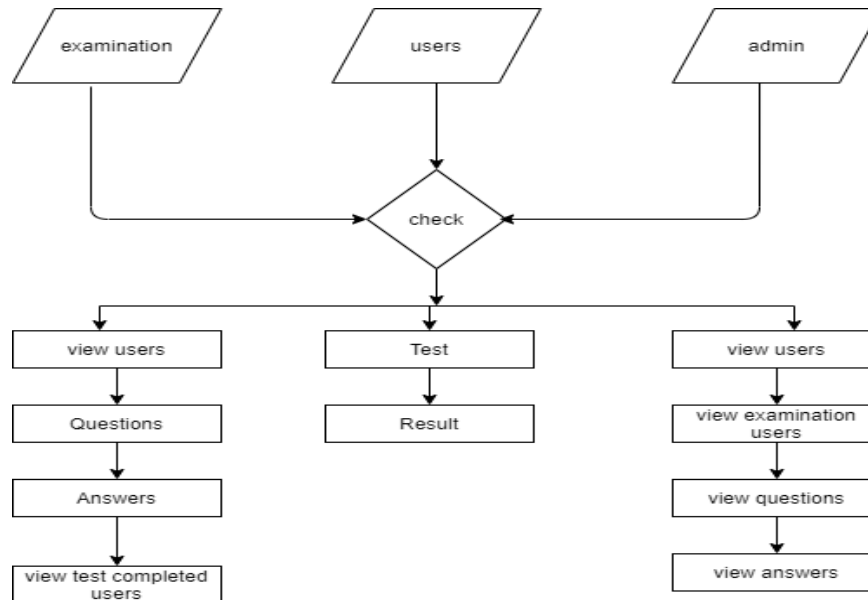
[3] SYSTEM ARCHITECTURE

Figure1. Data Flow Diagram For Online Examination

[4] SOFTWARE ENVIRONMENT

4.1 PYTHON

Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language. An interpreted language, Python has a design philosophy that emphasizes code readability (notably using whitespace indentation to delimit code blocks ratherthan curly brackets or keywords), and a syntax that allows programmers to express concepts in fewer lines of code than might be used in languages such as C++ or Java. It provides constructs that enable clear programming on both small and large scales. Python interpreters are available for many operating systems. C, Python, the reference implementation of Python, is open source software and has a community-based development model, as do nearly all of its variant implementations. C Python is managed by the non-profit Python Software Foundation. Python features a dynamic type system and automatic memory management. It supports multiple programming paradigms, including object- oriented, imperative, functional and procedural, and has a large and comprehensive standard library.

4.1 DJANGO

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app without needing to reinvent the wheel. It's free and open source. Django's primary goal is to ease the creation of complex, database-driven websites.
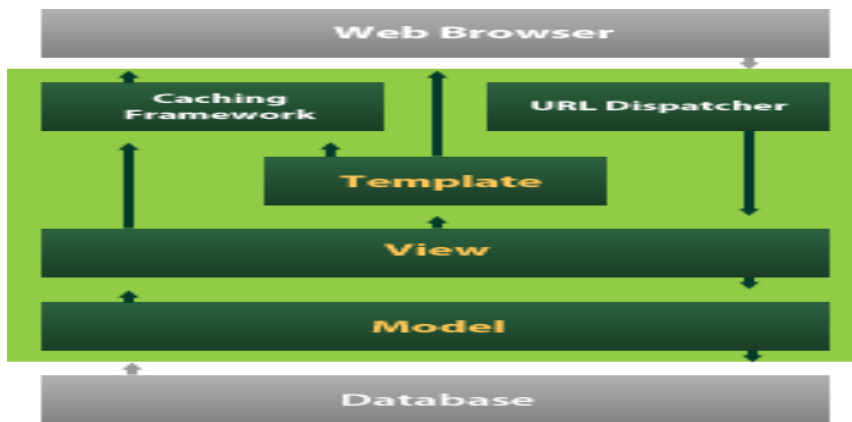
Django also provides an optional administrative create, read, update and delete interface that is generated dynamically through introspection and configured via admin models
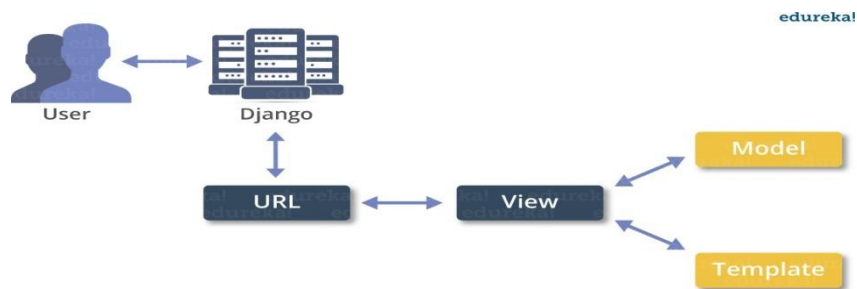
[5] IMPLEMENTATION

5.1 Modules Description

i) Exam cell: Examcell conducts the exams on particular technologies. first it will store all questionsabout particular technology into the database. Then store all answers of every technology into the database. After conducting the exam it will check the student answerswith database answers. Finally by using NLTK and Nlp concept give the result to students.

ii) Student: The content of it may be overlapping with that of others, but it also has its own discipline characteristics and is constantly developing new theories and methods. user after login into the session they will choose their particular exam .Then student will write the exam and submit the answers. after clicking test result automatically student will get the result based on their performance. all these concepts i.e getting of student score by using NLTK and NLP concepts.

iii) Admin: Admin will give authority to Exam-cell and students. In order to facilitate activate the Exam Cell and activate the students. the admin can see the details of Exam cell and students.. Admin can see the score of all students.

iv) Python and Data-Analysis: Python is an increasingly popular tool for data analysis. In recent years, a number of libraries have reached maturity, allowing R and Stata users to take advantage of the beauty, flexibility, and performance of Python without sacrificing the functionality these older programs have accumulated over the years. Python focus on simplicity and readability, python it boasts a gradual and relatively low learning curve. This ease of learning makes an ideal tool for beginning programmers. Python offers programmers the advantage of using fewer lines of code to accomplish tasks than one needs when using older languages.

[6] SCREEN SHOTS

Fig. 4 Home page

This is the Home Screen of Examination System, here we can see the home,examination cell, user and also admin.
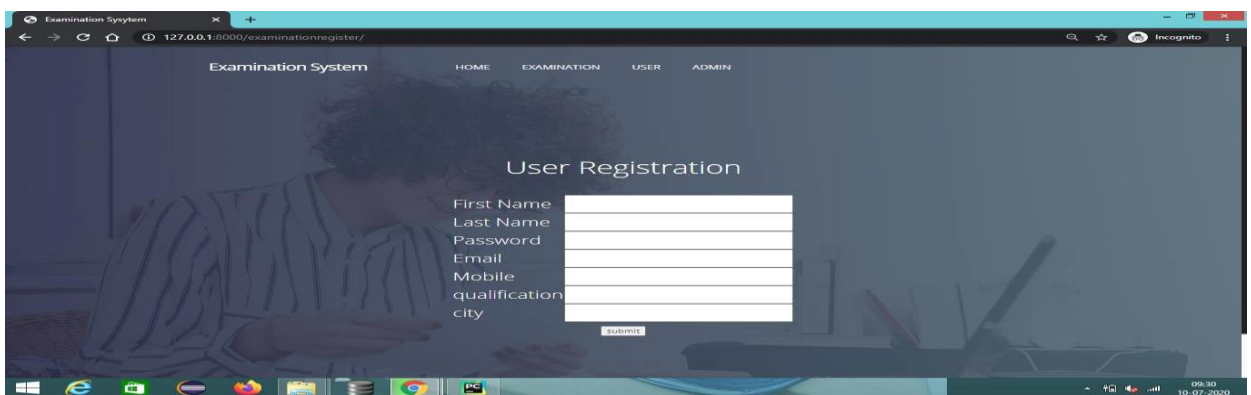

Fig. 5 Examination Registration

This Screen is an User Registration Page, Here we have to fill the form like first name,last name, password, email etc.
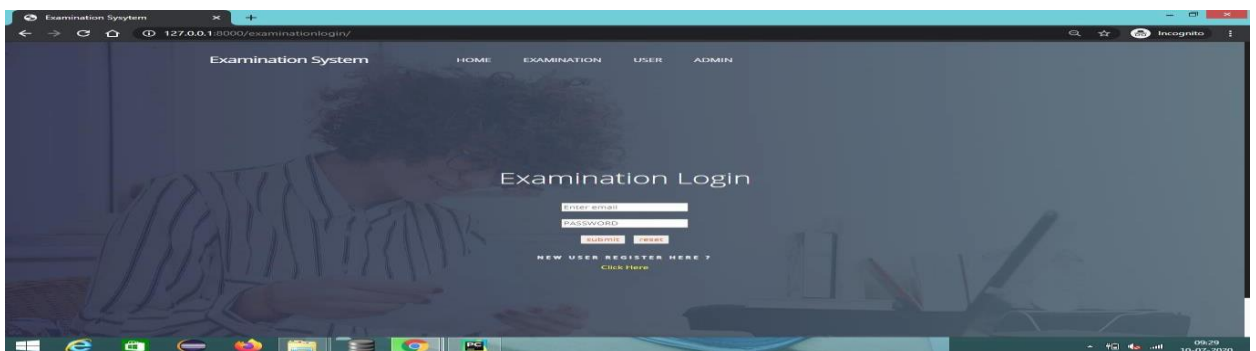

Fig. 6 Examination login

This image is an Exam Login Page, her we should enter exam mail and exam passwordthen we can login into the exam cell.

Fig. 7 Examination Home

This image is a Welcome Examination Page, when we login in to the Exam login then the welcome page will be appear in this page we can see the view Users, Questions, Answers, Test.
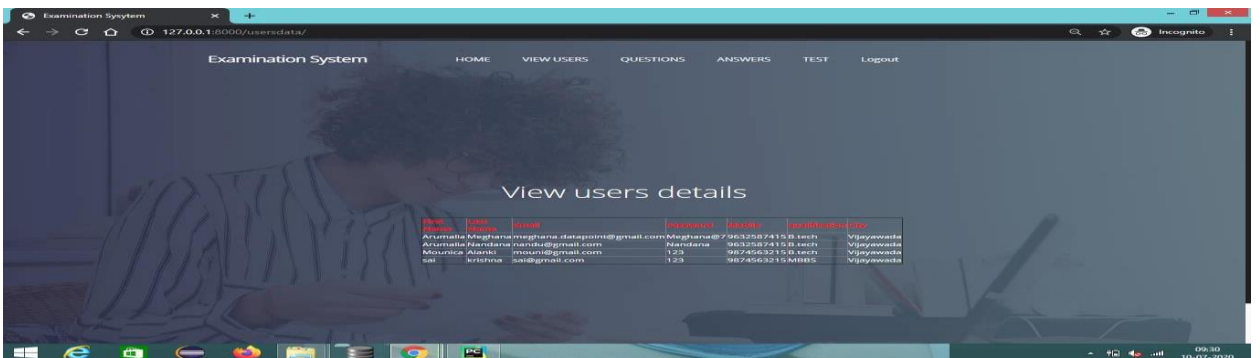
Fig. 8 View users

This image shows that View User Details Page, It consist of user details like first name,last name, email address, password, phone number, qualification, and also places.
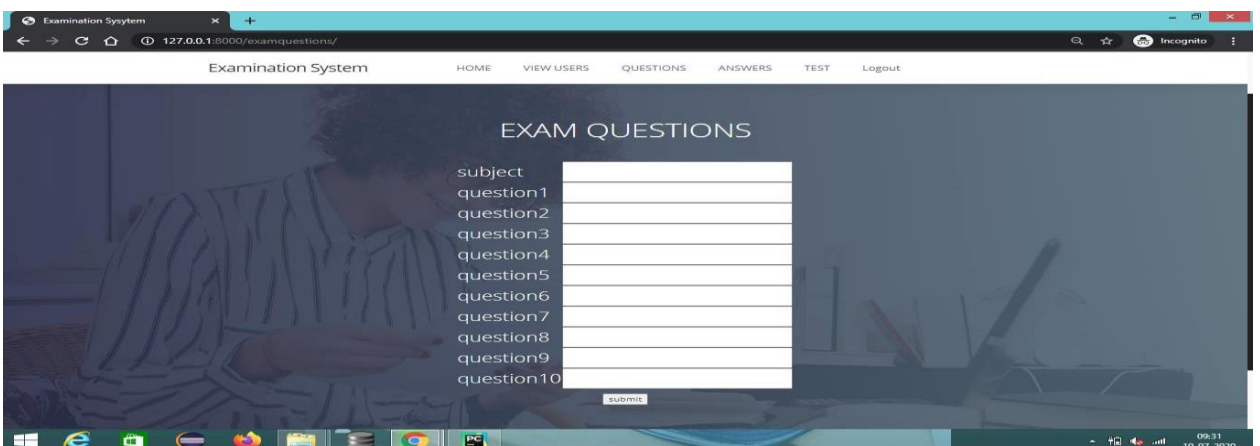
Fig. 9 Add Questions
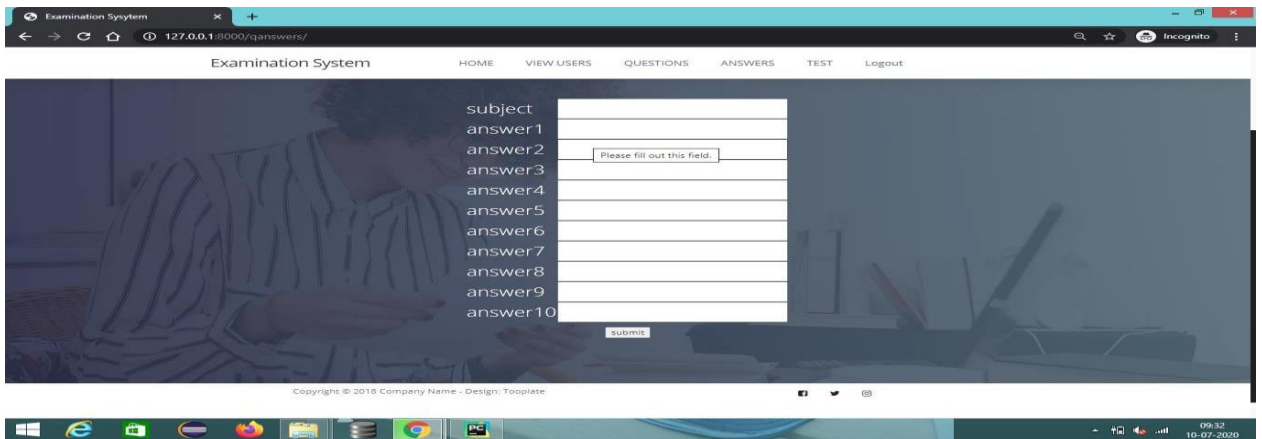
Fig. 10 Add Answers

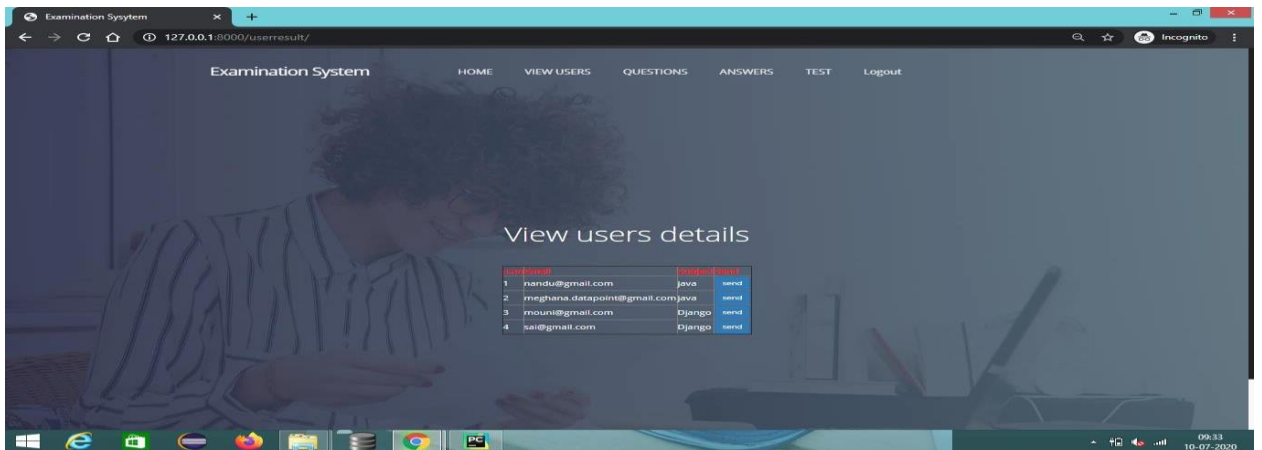This image is an Answer Page, in which we can write the answer for the given questions.

Fig. 11 Test Page

This Screen shot image is an view user details, by using this user details we send the test to the students.
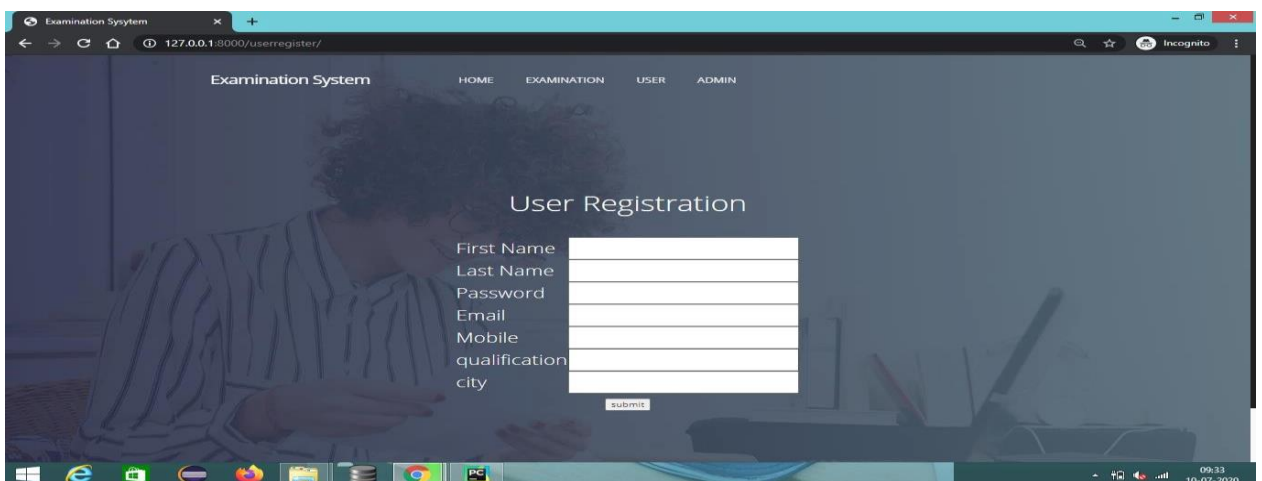

Fig. 12 User Registration

This image shows that User Registration, enter the first name, Last name, password,email, Mobile, qualification and city. We click on submit.
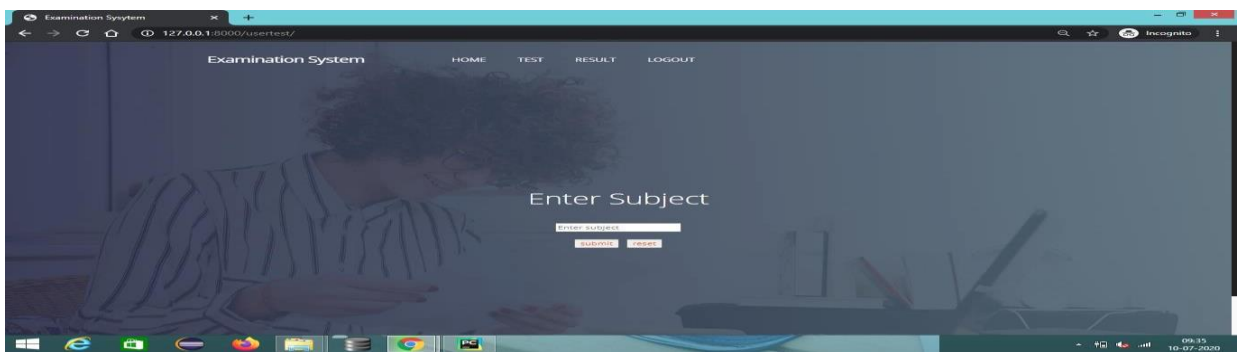
Fig. 13 Test Page

This image is an Subject Login, when we enter the subject name and we click on submit,then we can write the test of an subject.
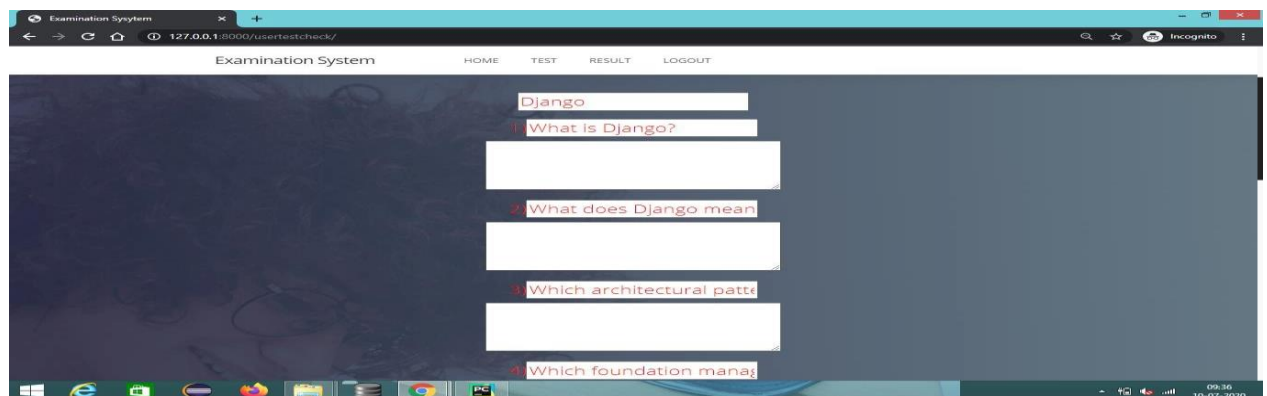

Fig. 14 Test Page

This image is an Test Page, in which it contains of "10 QUESTIONS" after we completion of the test ,we submit the test.
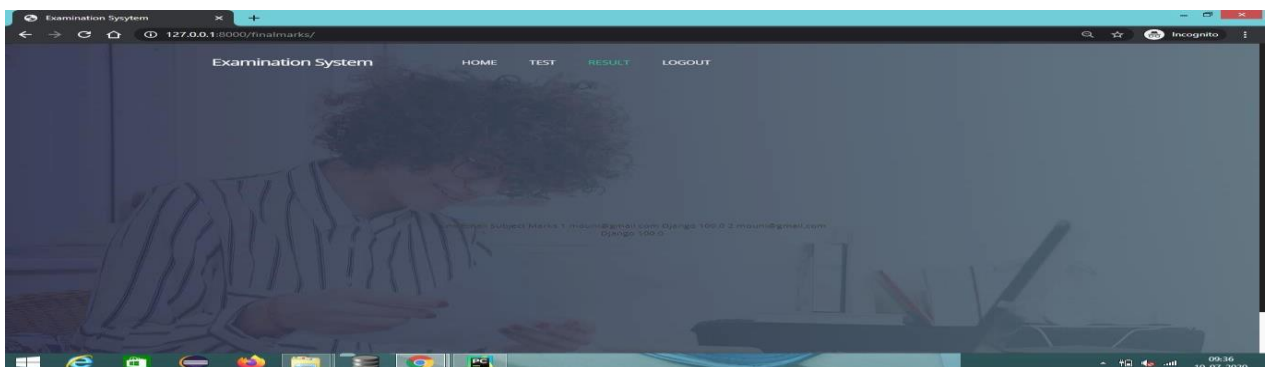

Fig. 15 Result Page

This image is an RESULT page, in this we can see our result of our exam in form of "Percentage".
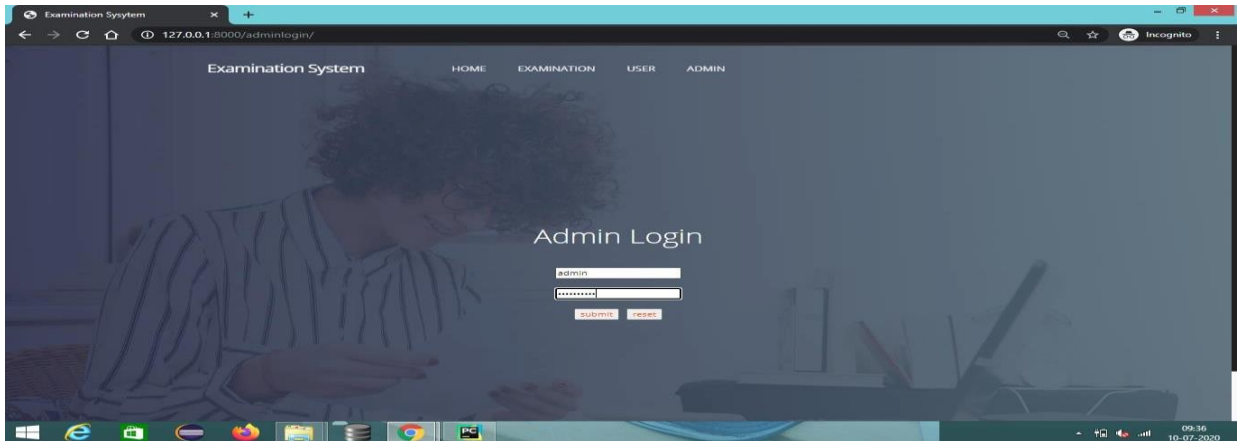
Fig. 16 Admin login

This image is an Admin Login page, when we enter the admin login, password and we clickon submit, then we login in to admin page.
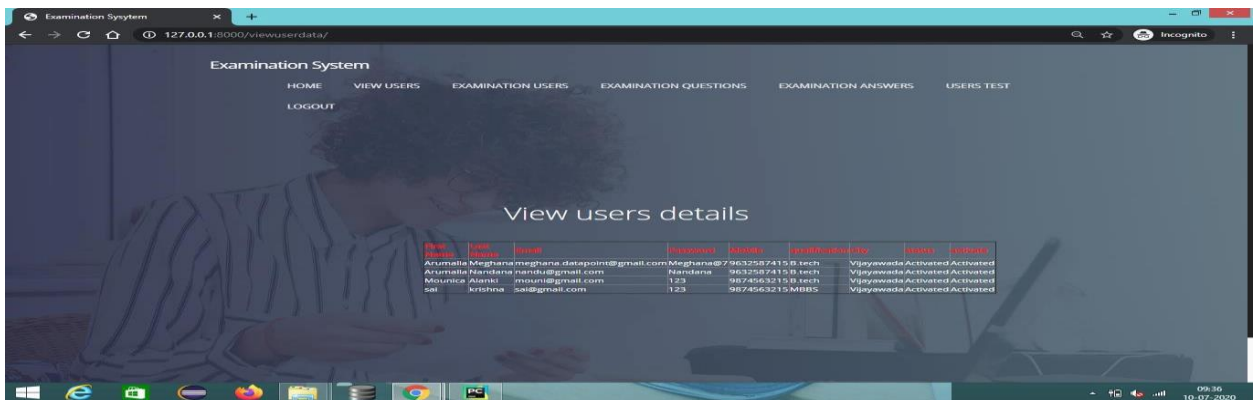

Fig. 17 View Users

In admin login we have View Users Details; in this we can see the user who are activated,Email, password, mobile number, qualification, first name, last name and also status.
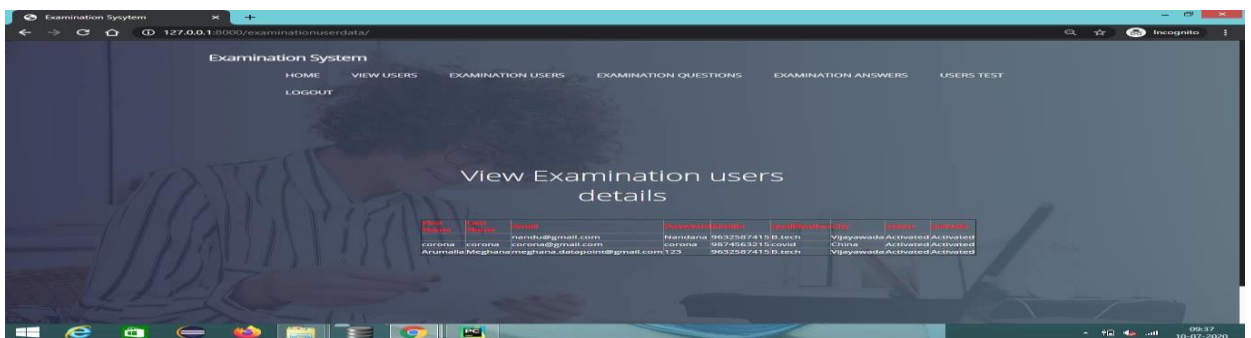

Fig.18 View Examination

In admin login we have View Examination Users Details, in this we have to activatethe user, then only user can write exam and submit the exam.

Fig. 19 View Questions Page

In admin login we consist of View Examination Users Details, It consist ofExamination question which are given by exam cell.
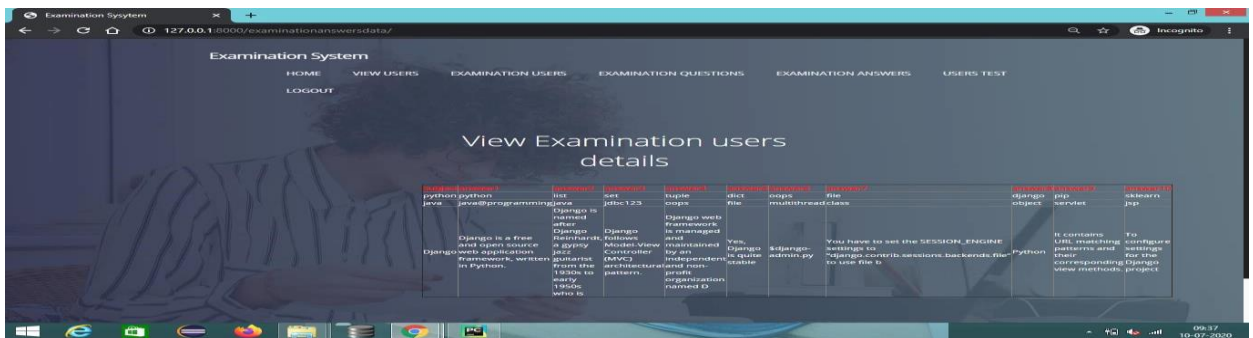

Fig. 20 View Answers

In admin login we have View Examination Users Details; it consists of answerswhich are given by exam cell.

[6] 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 theexisting 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 isrequired 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 thestudent can see the solutions and can correct the mistakes or errors committed while appearing for the exam.

The future works on it, it can be tracked using the system that which studenthas cheated from any other student. This will again use the concepts of Machine Learning and Data Science to work upon these.

REFERENCES

[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 Engineering, (2010).

[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).

[8] L. Zhai, and T. Gong "The research of examination management system based on network flat", IEEE 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), (2011).

[9] S. Luo, J. Hu and Z. Chen "Task Based Automatic Examination System for Sequenced Test", IEEE International Conference on Electronic Computer Technology, (2009).

[10] N. Yang, X. Chenguang G. Weiwei and M. Xianmin "The design of exam system on the basis of .net technology", IEEE Symposium on Robotics and Applications (ISRA), (2012).

[11] T. Treenantharath and P. Sutheebanjard "Secure Online Exams on Thin Client", IEEE 11th International Conference on ICT and Knowledge Engineering, (2013).

[12] Y. Atoum, L. Chen, A. X. Liu, S. D. H. Hsu, and X. Liu "Automated Online Exam Proctoring" , IEEE Transactions on Multimedia, (2017) .

[13] P M Yohan, Sk Althaf Hussain Basha, B Sasidhar, A.Govardhan , " Automatic Named Entity Identification and Classification using Heuristic Based Approach for Telugu", IJCSI(International Journal of Computer Science Issues), Volume 10, Issue 6,November 2013,ISSN: 1694-0814.

[14] Naga Raju Devarakonda, SK Althaf Hussain Basha, ShaikSubhani,"Outliers Detection in Regression Analysis using Partial Least Square Approach", ICT and Critical Infrastructure: proceedings of the 48th Annual Convention of Computer Society of India- Springer, Vol II Advances in Intelligent Systems and Computing, Volume 249, pp. 125-135, Visakhapatnam, December 2013,ISBN: 978-3-319-03095-1.

[15] Jinka Sreedhar SK Althaf Hussain Basha, Pammi Pavan Kumar, "Innovative Techniques and Technologies in Translation in a Multilingual Context -2012", Third International Conference on Translation, Technology and Globalization in Multilingual Context, ITA, NewDelhi

[16] Y.SriLalitha, Sk Althaf Hussain Basha, N.Sandhya, A.Govardhan, "Web Usage Mining- Survey ", Published in proceeding of National Conference on "Soft Computing and Artificial Intelligence" (NCSCAI 2009) at Lingaya's University,Faridabad,222-227,2009.

[17] SkAlthaf Hussain Basha, Jinka Sreedhar, "Introduction: Behaviour of Cuckoo Algorithm", Advanced Engineering Research and Applications Vol. VIII (Book Code: AERA), Research India Publications,978-93-87374-32-4,2018.

[18] Sk Althaf Hussain Basha, Jinka Sreedhar, Introduction: Behaviour of BAT Algorithm", Advanced Engineering Research and Applications Vol. VIII (Book Code: AERA), Research India Publications,978-93-87374-32-4,2018.

[19] Jinka Sreedhar, Sk Althaf Hussain Basha, "A Case Study : t - Closeness through Micro aggregation: Strict Privacy with Enhanced Utility Preservation in Sensitive Microdata", Hand book on Computer and Information Technology Vol.II (Book Code: HCIT),Research India Publications, ISBN: 978-93-87374-25-6,2018.

[20] SkAlthaf Hussain Basha, Jinka Sreedhar, Implementation: Ranking and Suggesting Widespread Items In E-Commerce Applications ", Handbook on Computer and Information Technology Vol. II (Book Code: HCIT), Research India Publications, ISBN: 978-93-87374- 25-6,2018.

[21] Dr. G. N. R. PRASAD, "Identification of Bloom's Taxonomy level for the given Question paper using NLP Tokenization technique", Turkish Journal of Computer and Mathematics Education, Vol.12 No.13 (2021), 1872-1875.