# major_project_research_paper.docx

*by*

**Submission date:** 28-Apr-2022 02:38PM (UTC+0530)
**Submission ID:** 1822689914
**File name:** major_project_research_paper.docx (2.21M)
**Word count:** 2095
**Character count:** 11105

# 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 various levels and even a few portals are created to conduct school or college examinations, along with placement type of exams however the portals in general, are designed for the same behaviour that contains a few choice questions and the examinee should choose one option from it. The main aim to not build the technology that is currently available, but a very different project that solves the big problem of online testing. Here we use a descriptive online test program. Single choice questions or multiple-choice questions are easy to conduct and test with as it will have a question, and along with four different options, where we have to choose one or correct answer with respect to the question along with the particular discipline of subject or particular topic within the equal question that keeps the right choice within the website. However, this is not the case with the descriptive questions, With the help of 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 system of current online examination portals has multiple choice type questions, although most exams have detailed questions, which have multiple select answers. They are not efficient in testing the knowledge of the students and, therefore, it will not be convenient nor effective to replace them on a large scale. We all know that if we want to get rid of the commonly accepted system of testing the students and how in general the tests are conducted, the new system must not only be best in terms of quality and user experience, but also change in quality of the questions and the correct proctoring and security of the examination

systems so that the exams can be conducted in the highest trust and integrity possible with the trust of the companies.

**Problem Description**

The current model which we are building, we want to address the main problem of online system portal by building the feature where the student writes the descriptive answers which will be automatically evaluated with the machine learning model which we will build using natural language processing techniques with the help of computing efficiency without any human error. This can be done using Text analysis and Semantic matching using various NLP techniques such as tokenization, Lemmatization, stemming, TF-IDF and BERT algorithm. The answers written by the students will be evaluated and stored in the portal, where the student can login and check the marks scored for a particular subject.

**Purpose of the Project**

The current project which we built will give a big jump to online examination system how they function as it allows it to overcome its main con and helps the online examination system to raise its features and be more relevant for annual examinations conducted for school or college evaluations. This system frees teachers and professors from the burden of checking copies and where they can invest more time in teaching, which will eliminate bias reviews 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 wastage of resources used in conducting the examinations, which will help in resource management as it is where the cost of examination increases. 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 collection and processing, is the first step of this project where the data will be generated by the examination, where the faculty will generate questions and answers for the particular subject, the questions and answers then are collected the pre processed for the further steps to model training

Then the next step is tokenization , second step in this the the entire the sentences are divided into strings then comes the third step where lemmitization on this strings is done along with stemming where each word is given a category to which it belongs then comes the next step of Keyword extraction by using the algorithm TF-IDF which gives the ranking for each keyword based on the occurrence of that string for evaluating answers for questions that are descriptive in nature. Keywords are stored for every question and on the basis their occurrence, each word is given the ranking, for the final step comes with matching of the answers written by the teacher and student by using BERT algorithm which will use the bi-directional conditions to evaluate the answer strings by both faculty and student and then with the model built it will match the strings and allots the marks to the students for the answers written.

Speaking of the technologies used to create such a model to test descriptive responses, NLP or Natural Language Processing has a major role to play. For those of you who don't know about the machine behind this, it's all about playing with a series of letters, numbers and special characters or what we call string tricks to get those results.

**Existing System**

This world has seen many test portals used on multiple servers used to conduct online tests for various Olympiad purposes and some i.e., designed to perform experiments for placement purposes. But what we have seen is that most portfolios are designed to do tests that contain multiple choice questions.

**Disadvantages**

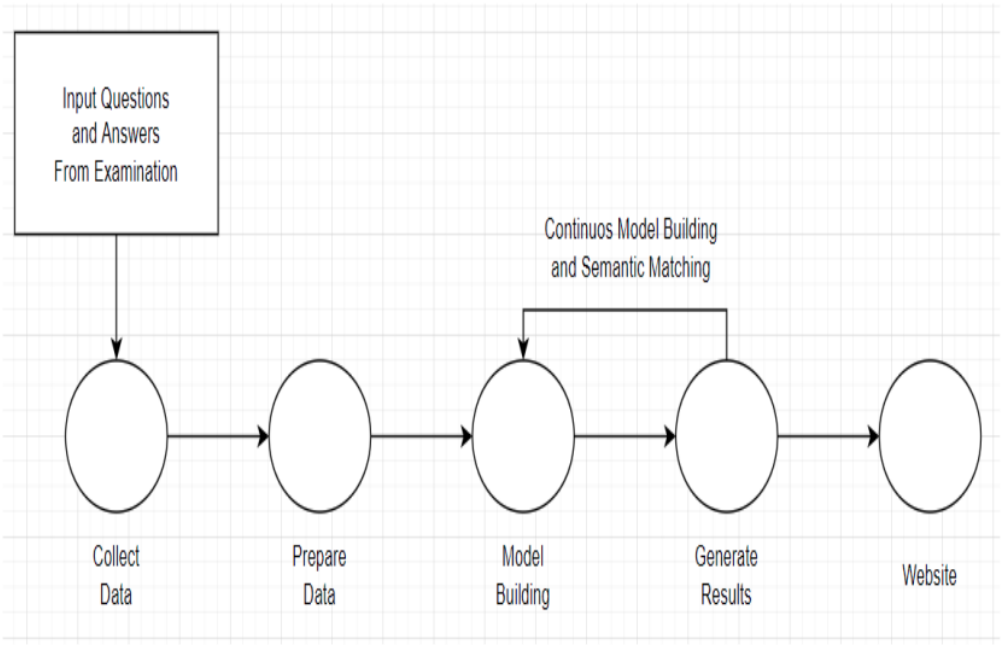
The offline test system is very costly and not so efficient due to the resources used and the time spent and the existing online test system includes only a variety of multiple-choice questions.

**Proposed System**

The proposed model takes the online testing system to another level by enabling the test subject to write descriptive responses that will be automatically tested. Tested responses will be stored on the website and can be accessed at any time and a special student profile will be stored for better student testing. Speaking of the method used to construct such models to test descriptive responses, NLP or Natural Language Processing has a major role to play. The online test system works very well and the answers are tested on their own and the student can see the solutions and can correct any mistakes or mistakes made while registering 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.

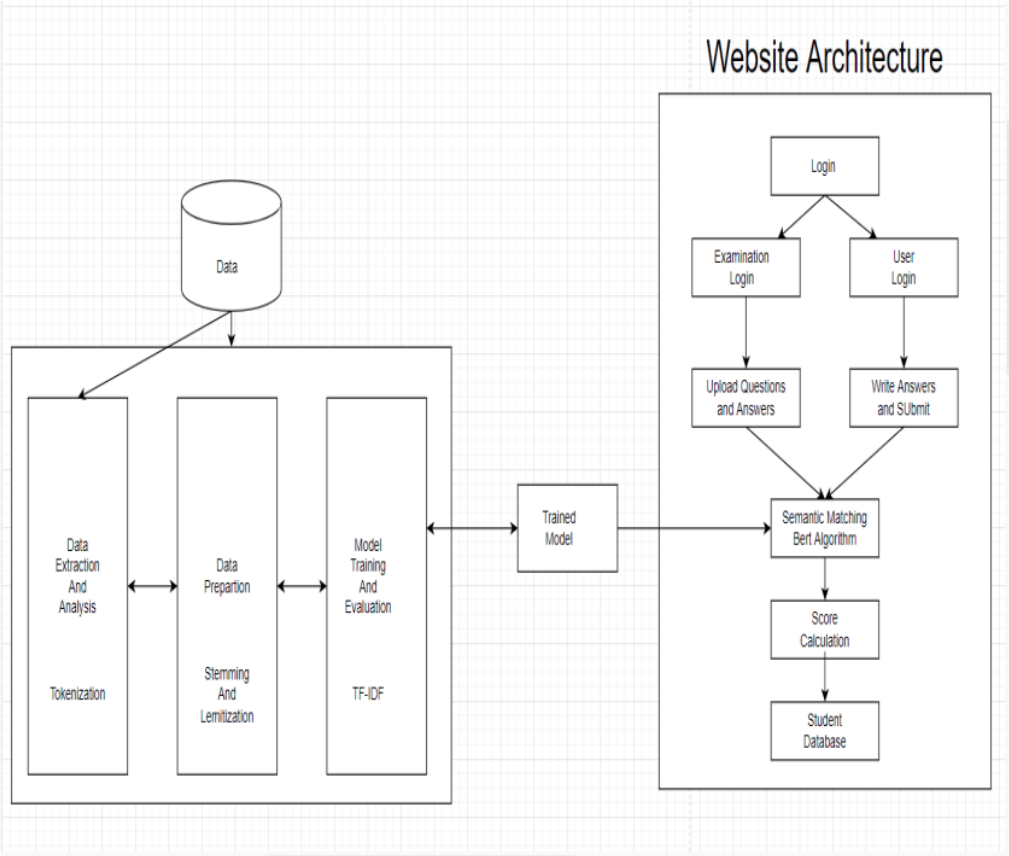



Collecting data is the first step of the process flow which includes defining project, setting up the machine environment suitable for the development requirements and later understanding the data

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.

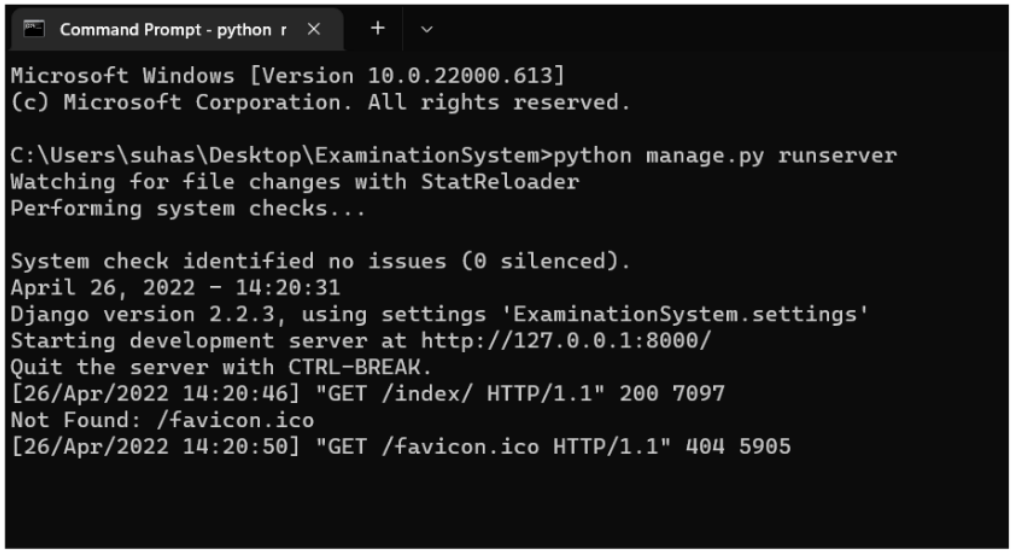
**RUNNING THE MODEL**

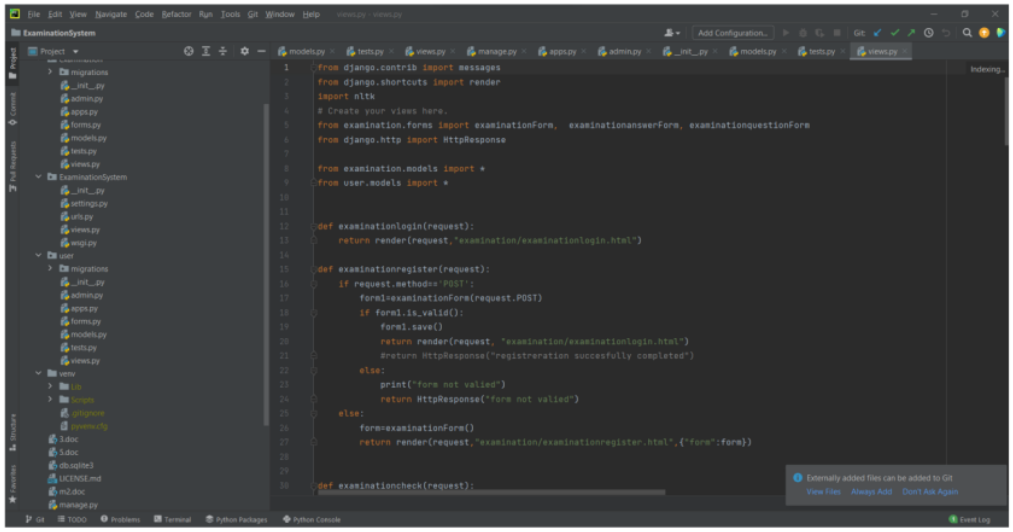
3

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

4

- 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 of the examination portal

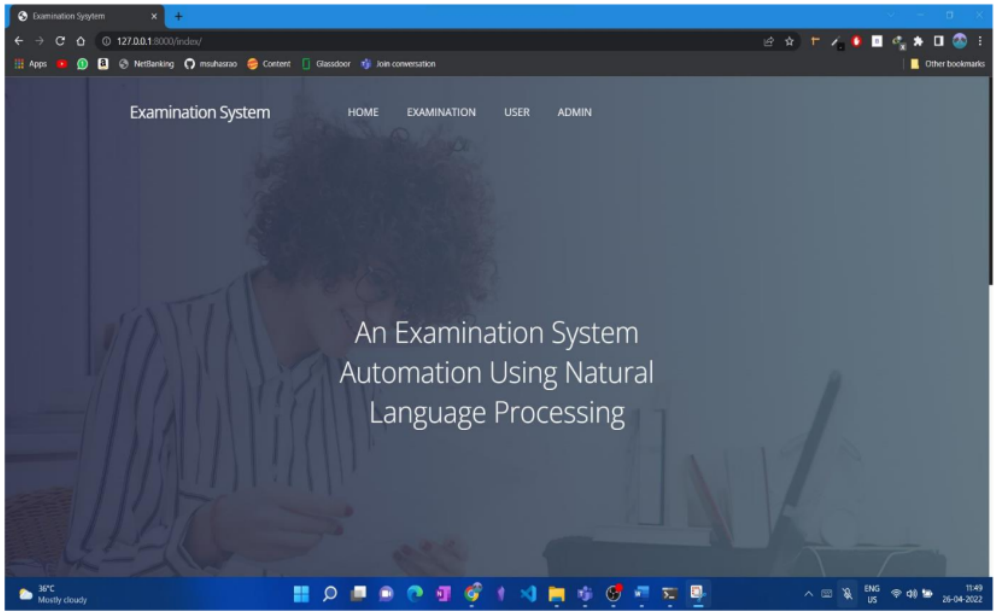
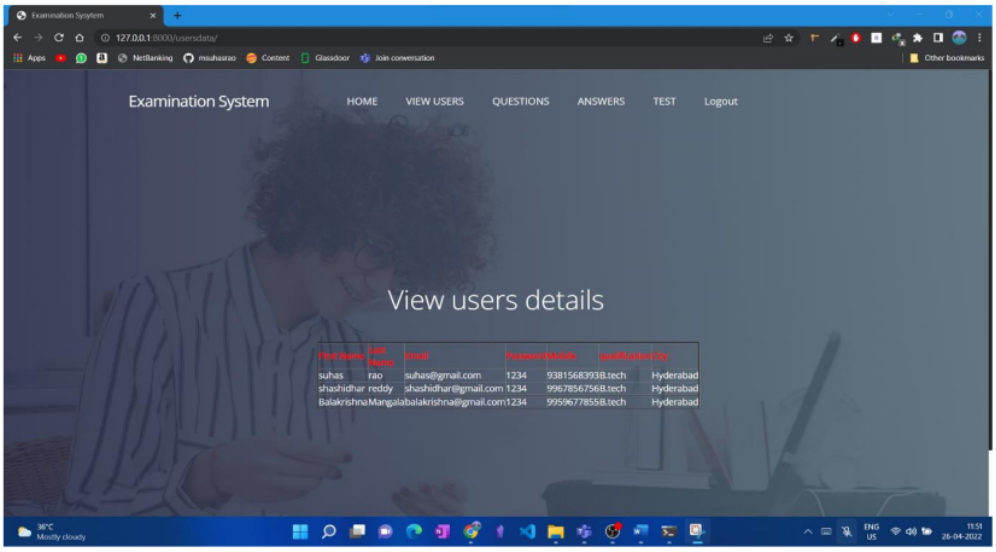



**Fig Home page**

**It consists of 3 tabs examination, user and admin**



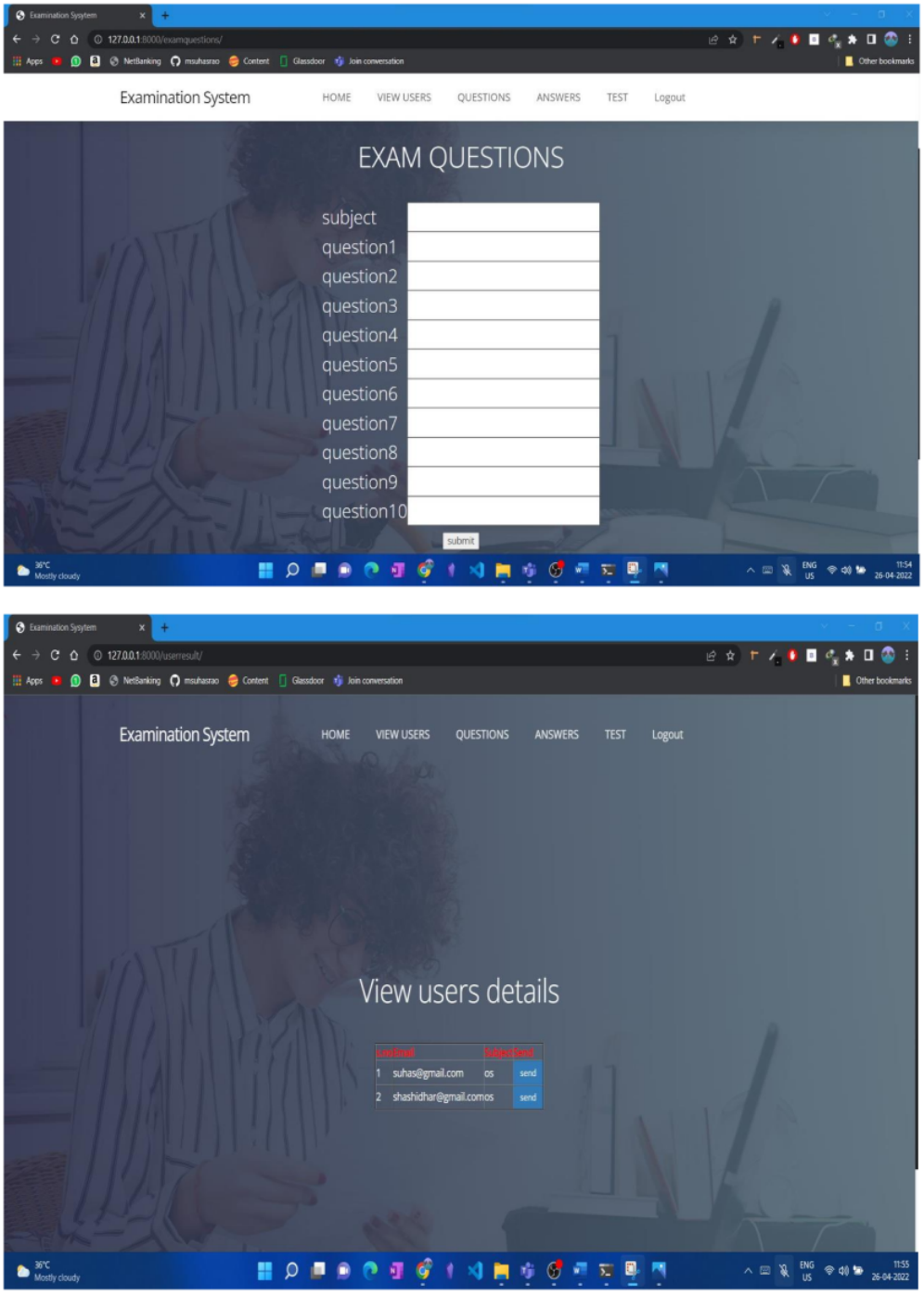



**Fig Test result page**

## CONCLUSION

In conclusion, it can be seen by doing experiments using such an algorithm from time to time that one can determine the tendency of the marks obtained by different students and we can give them an analyzed report on the various topics they need to focus on. the weak. With the available data, we can also use a data learning machine model that predicts data that students will receive in the future. It is noteworthy that students learn mainly those subjects that are intended or required for the sole purpose of inclusion. While students become subjects of their core background. In-depth domain information is required because it is not helpful to read if you do not have background information. So it can help students to get quality information as everything will be digital and there will be no complicated process of doing paper tests. Also, the answers are checked quickly and the student can see the solutions and can correct any mistakes or mistakes made while appearing on the test. a social networking site and online forums from various sources that host different companies. The Online Examination System is widely used compared to other tests. These bugs need to be identified and resolved in order to improve software quality. Therefore, in the future we can develop more secure software using advanced technology, by introducing new security systems using biometrics, we can identify student identity by analyzing digital signature or fingerprints or photography and using live internet 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)

# major_project_research_paper.docx

ORIGINALITY REPORT

| 17% | 4% | 15% | 0% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | Indrashis Das, Bharat Sharma, Siddharth S. Rautaray, Manjusha Pandey. "An Examination System Automation Using Natural Language Processing", 2019 International Conference on Communication and Electronics Systems (ICCES), 2019<br>Publication | 13% |
| 2 | journals.pen2print.org<br>Internet Source | 3% |
| 3 | git.its.aau.dk<br>Internet Source | 1% |
| 4 | www.learnpy.in<br>Internet Source | 1% |

| | | | |
|---|---|---|---|
| Exclude quotes | On | Exclude matches | < 10 words |
| Exclude bibliography | On | | |