Final Year Project Presentation

Guided By : Ms. Mudita Sharan Ma'am

Supervised By : Mr. Pawan Kumar Tiwari Sir

Problem Statement

Conducting examination and answer sheet evaluation are hectic testing tools for assessing academic achievement, integration of ideas and ability to recall, but are expensive, resource and time consuming to generate question and evaluate response manually.

Manual evaluating of answer sheet takes up a significant amount of instructors' valuable time and hence is an expensive process.Also different security concerns regarding paper leakage is one of the other challenges to conquer.

Ever thought how hectic it sometimes becomes for the teacher to select good questions from the set of defined syllabus? Even what’s more hectic is the evaluation of the answer sheets and display of the marks to the students.

Solution : Testana (A examination Tool)

01 We made a web-application powered by latest technology so that the teacher who faced the problem of creating questions and evaluating can use it to make more hassle free.

02 After the questions are being answered we are using NLP to evaluate the questions answer. This helps the teacher to get rid of the correction process which takes a good amount of time. WE have made that possible in few seconds.

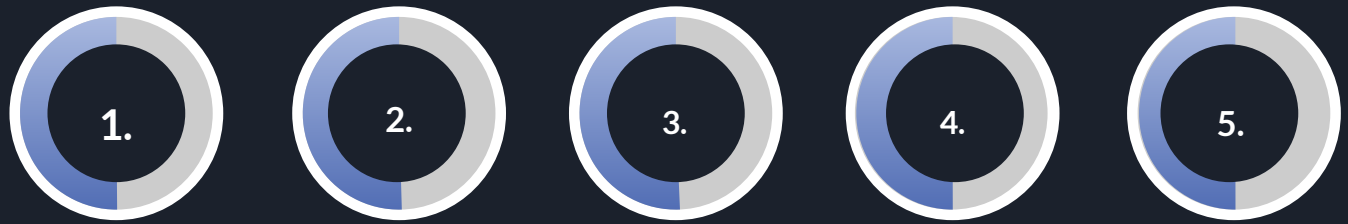
03 After the evaluation is done, the marks display is also a time taking process. We in the end have automated the structure in which the marks generated by the NLP using Cosine SImilarity is stored in the excel sheet directly.

Project objective

TESTANA aims to build an automated examination system using machine learning, natural language toolkit (NLTK), python environment, flask framework, and web technologies to provide an inexpensive alternative to the current examination system. We implement a model to automatically generate questions with their respective answers and assess user responses.

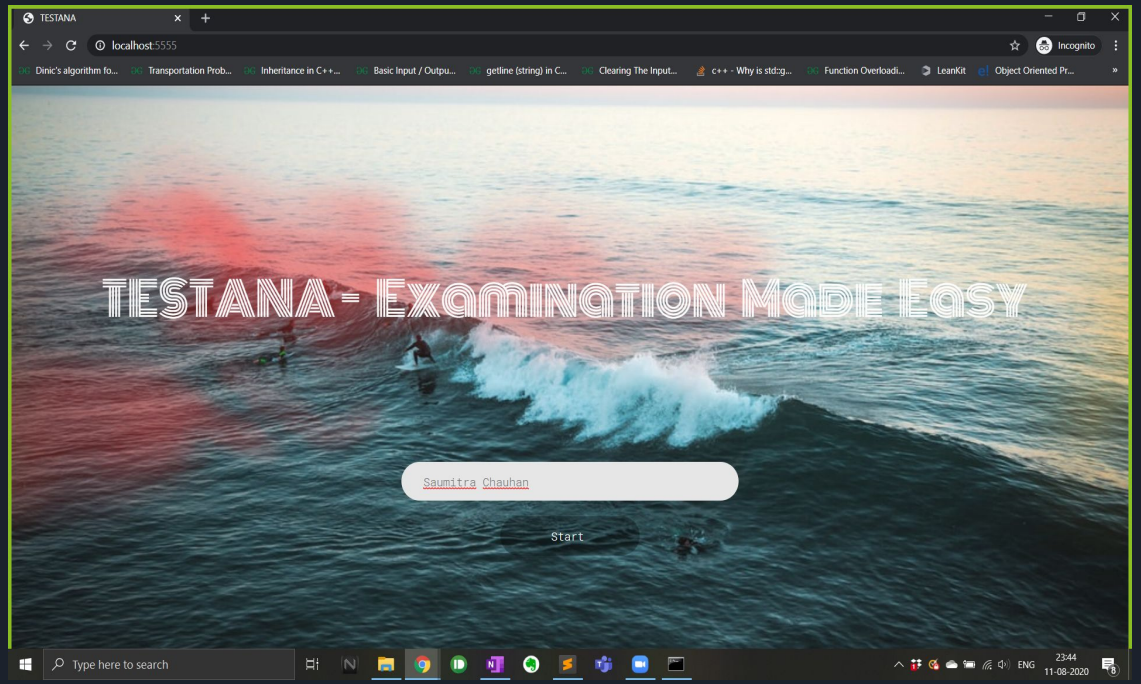
Framework of TESTANA

We are building our application to be user-friendly usage and anyone with little knowledge of computers also be able to use our application. The user will login to our web app and a session will be created for that user. The application will have different pages where the subjects will be shown on the page and the user can choose from the desired subject list.



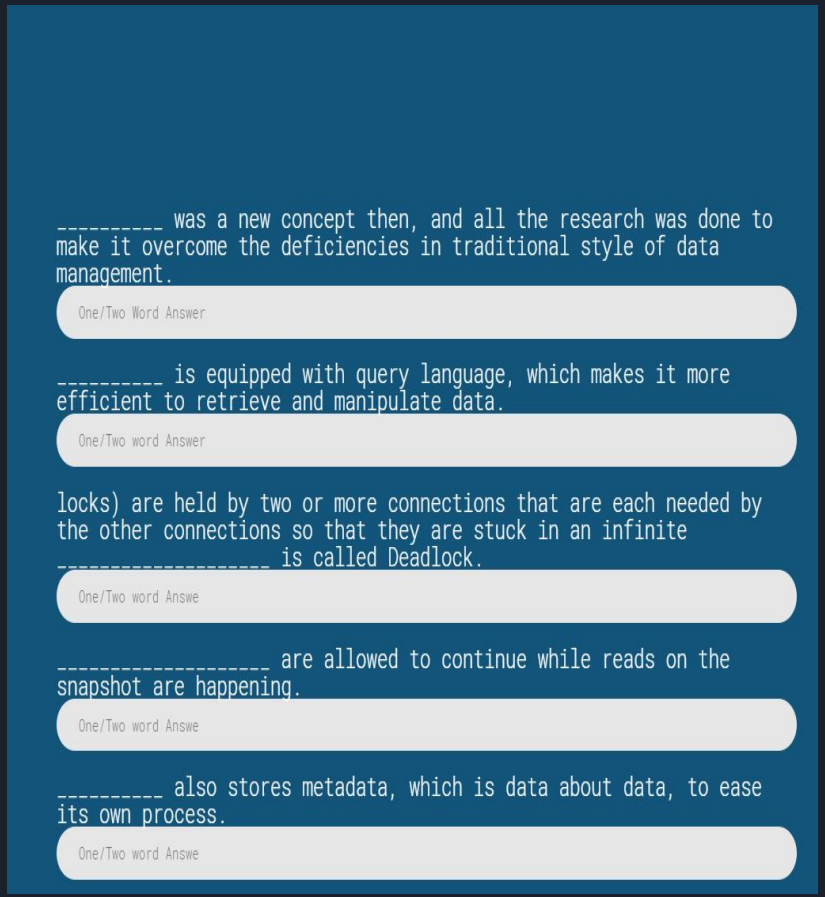
The Web Application

The user will login to our web app and a session will be created for that user. The application will have different pages where the subjects will be shown on the page and the user can choose from the desired subject list. It will also have a page 2 different pages for subjective and objective questions respectively. We will be testing the web app for 10 questions on the objective page and 10 questions on the subjective page.After that an evaluation page will be also there.



Generating Random Questions and Their Respective Answers

In order to generate questions and answers, the text content will be put into the process of Information Extraction. Information Extraction is the task of automatically extracting structured information from unstructured and/or semi-structured documents.



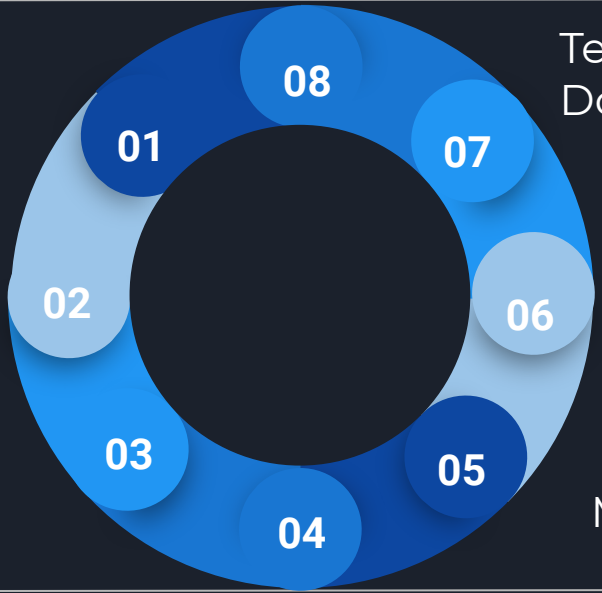
Steps involved in Question Generation & Answer Generation

01 Tokenization

02 Stop Word Removal

03 Stemming

04 Chunking



05 Marking Unmovable Phrases

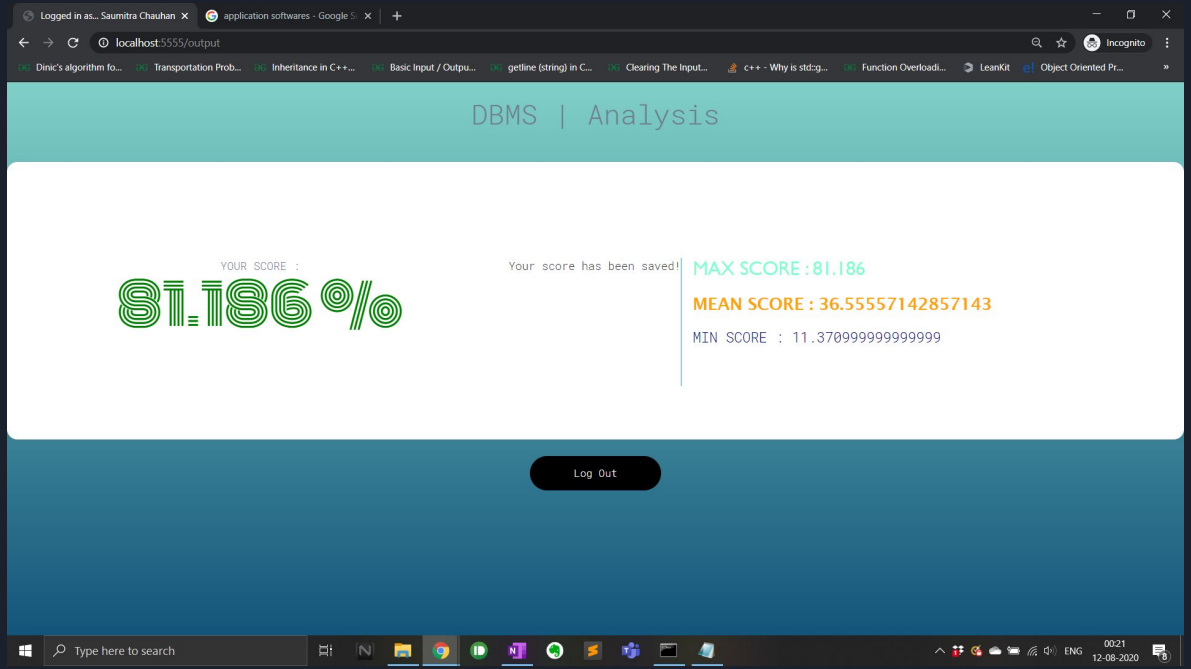
06 Generating Question Phrases

07 Generating Answer Phrases

08 Term Frequency – Inverse Document Frequency (TF-IDF)

User Input & Evaluation

The user answers will be brought and firstly preprocessed using these methods for later evaluation. Then the score would be computed on matching the student answer and the system generated answer using **Cosine Similarity** algorithm and give a particular score on the basis of the degree of match.



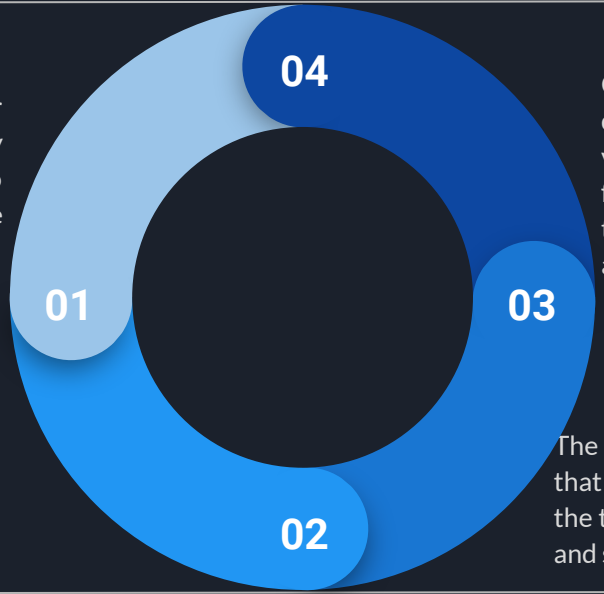
Steps involved in User Input and Evaluation

Tokenization

The big chunks of text which are smaller than large strings can be processed by tokenization and converted into sentences, then at the end these sentences can be tokenized into words.

Stop Word Removal

This step can improve the performance of the system by reducing the size of text by 20-30% of overall word counts in a specific text document



Cosine Similarity

Cosine similarity is used to measure the degree of similarity between two vectors. Weighted term results are used for the calculation of similarity between the reference answers and the student answers.

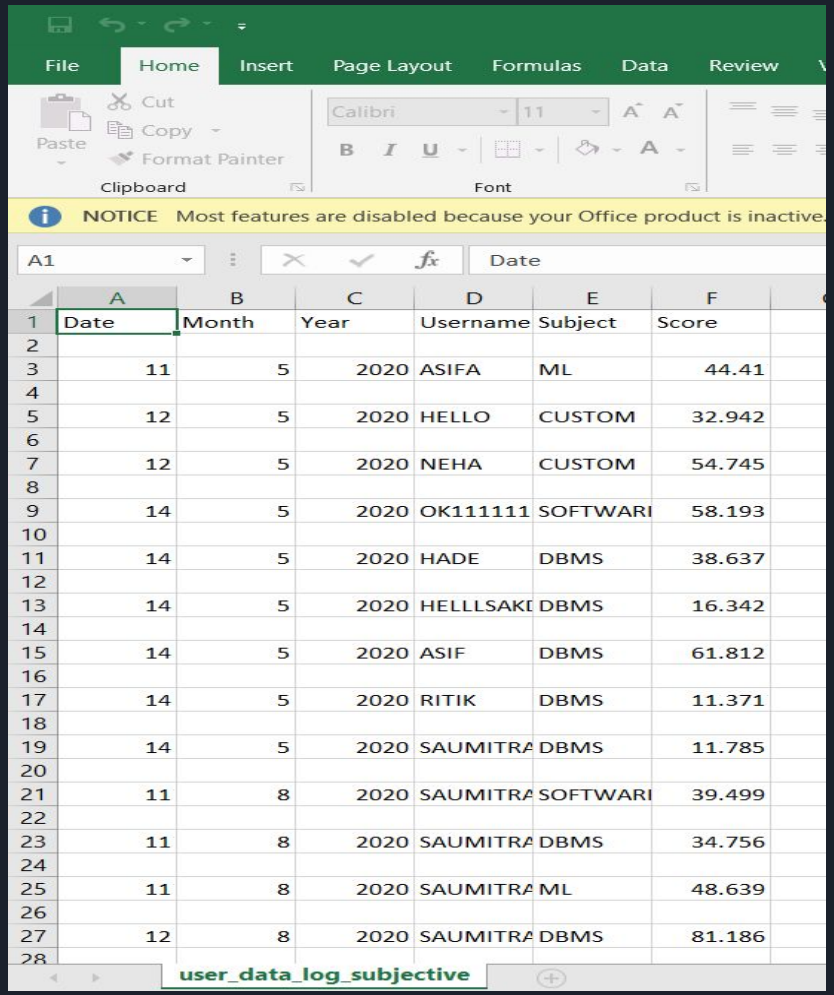
Stemming & Term Frequency

The purpose for stemming process is the words that sharing the exact root or stem will reduce the total number of words that need to process and save time and memory space.

Storing of the Result

The result would be stored in the form of the table in the excel data sheet. The following data would be saved for each user:-

1. UserName
2. Subject Name
3. Marks Scored
4. Date of Examination



File Home Insert Page Layout Formulas Data Review

Cut Copy Format Painter Paste Clipboard

Calibri 11 Font

NOTICE Most features are disabled because your Office product is inactive.

A1 Date

	A	B	C	D	E	F
1	Date	Month	Year	Username	Subject	Score
2						
3	11	5	2020	ASIFA	ML	44.41
4						
5	12	5	2020	HELLO	CUSTOM	32.942
6						
7	12	5	2020	NEHA	CUSTOM	54.745
8						
9	14	5	2020	OK111111	SOFTWARI	58.193
10						
11	14	5	2020	HADE	DBMS	38.637
12						
13	14	5	2020	HELLLSAKI	DBMS	16.342
14						
15	14	5	2020	ASIF	DBMS	61.812
16						
17	14	5	2020	RITIK	DBMS	11.371
18						
19	14	5	2020	SAUMITRA	DBMS	11.785
20						
21	11	8	2020	SAUMITRA	SOFTWARI	39.499
22						
23	11	8	2020	SAUMITRA	DBMS	34.756
24						
25	11	8	2020	SAUMITRA	ML	48.639
26						
27	12	8	2020	SAUMITRA	DBMS	81.186
28						

user_data_log_subjective

Technology used in TESTANA

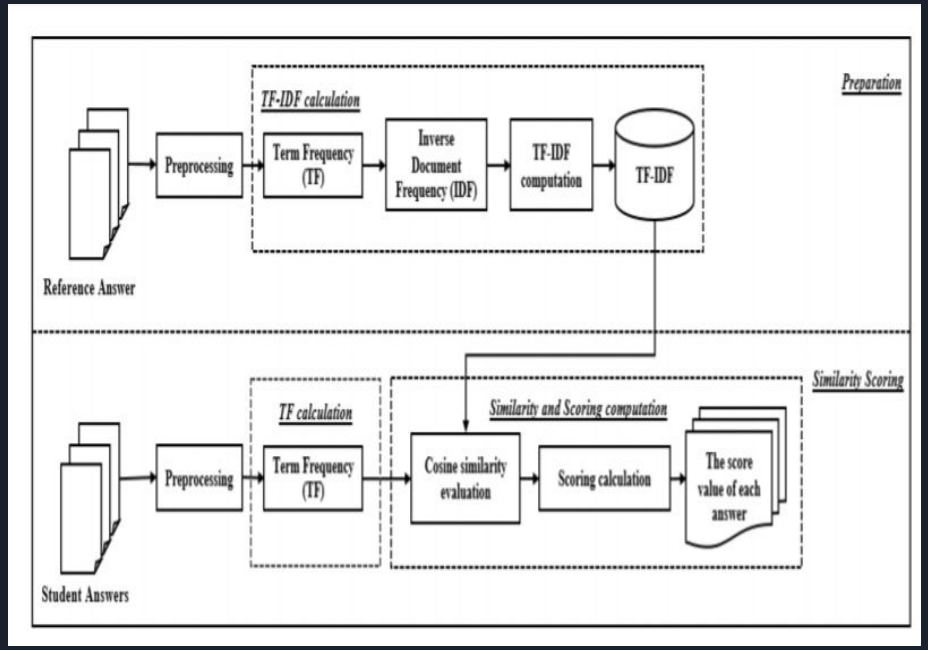
TESTana will use these above technologies to implement the model and web technologies such as HTML, CSS, Bootstrap, JavaScript. Python v3.6 is the main backend language we will be using to execute the desired result.

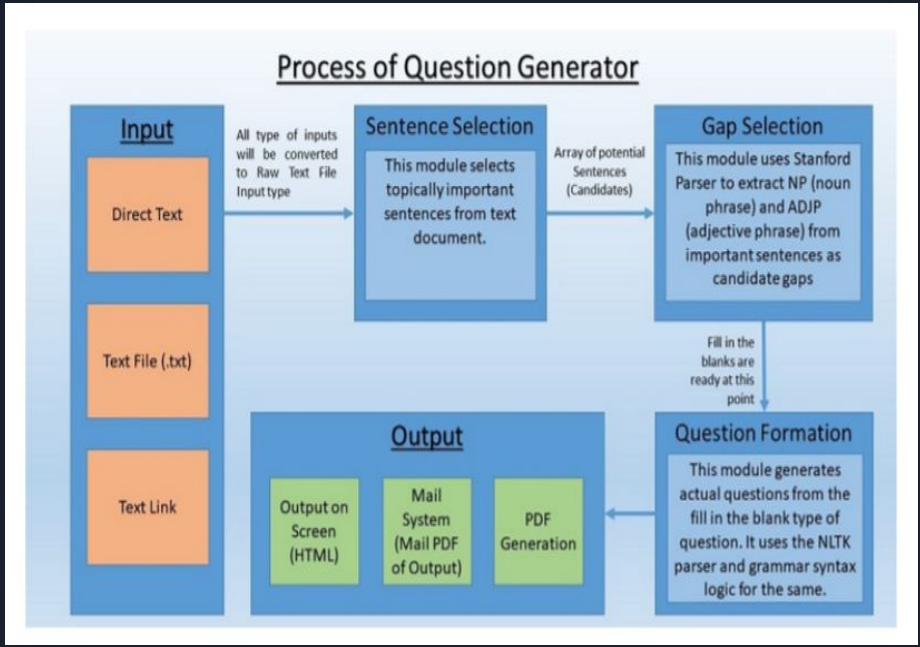
1. Flask 0.12
2. Numpy 0.10
3. Pandas 0.20
4. TextBlob 0.10
5. NLTK 3.3

Process of the Project

In this we present the computerized examination system, which is divided into two main stages: the preparation stage and the similarity scoring stage. First, the preparation stage aims to prepare a reference answer in terms of the TF-IDF vectors, consisting of two processes; preprocessing and TF-IDF calculation. Secondly, the similarity scoring stage aims to evaluate the student answers, and determine the student scores. There are three processes in this stage; preprocessing, TF calculation, and similarity and scoring computation.

Process Charts :





Software Requirements

1. HTML5 - Use to design web pages.
2. CSS3 - Used to style web pages
3. Bootstrap 4 - Make web app as mobile friendly
4. JavaScript - Helps in creating web pages
5. Flask - Frontend Framework
6. Python 3.6 - Backend Language
7. NLTK Library - Helps in NLP like Tokenization
8. TextBlob Library - Helps in NLP
9. Pandas and Numpy Libraries - Python libraries
10. MS-Excel - Database

Result & Conclusion

In order to evaluate the performance of the proposed system, we conducted the following experiment: We took data of the “Database Management System and Database Design” course to prepare a question dataset and asked our user to answer each question. We then asked 10 friends to join the experiment by giving their answers to the questions. The mentor then evaluated the answers, and provided a score for their answers, which were then compared with scores obtained from the proposed system in both methods; namely with and without the synonym, in which to compare the similarity with the reference answer.

In study we developed a tool to calculate the results using Natural Language Processing (NLP). The result is obtained through the NLP algorithm and the result is produced as a final result. These algorithms are used successfully and produce efficient results. The system generated results are compared with evaluation done by the faculty members. We found that there is no vast difference between the system generated marks and evaluation done by the faculty member.

Introducing: Team

It would not have been possible with this team and our mentor.

Saumitra Chauhan
Backend Developer

Ritik Rajak
UX Designer

Asif Ali Khan
Front-End Developer

THANK YOU