

**INTELLIGENT SYSTEM FOR DESCRIPTIVE ANSWER
EVALUATION**

A PROJECT REPORT

Submitted by

SHAD IBRAHIM N P (TKM20MCA-2034)

to

The APJ Abdul Kalam Technological University

In partial fulfillment of the requirements for the award of the Degree of

MASTER OF COMPUTER APPLICATIONS



**Thangal Kunju Musaliar College of Engineering
Kerala**

DEPARTMENT OF COMPUTER APPLICATIONS

JULY 2022

DECLARATION

I undersigned hereby declare that the project report INTELLIGENT SYSTEM FOR DESCRIPTIVE ANSWER EVALUATION, submitted for partial fulfillment of the requirements for the award of degree of Master of Computer Applications of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by me under supervision of Prof.NATHEERA BEEVI M . This submission represents my ideas in my own words and where ideas or words of others have been included, I have adequately and accurately cited and referenced the original sources. I also declare that I have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Place: Kollam

Date: 22-07-22



SHAD IBRAHIM N P

DEPARTMENT OF COMPUTER APPLICATIONS
TKM COLLEGE OF ENGINEERING



C E R T I F I C A T E

This is to certify that, the project report entitled **“INTELLIGENT SYSTEM FOR DESCRIPTIVE ANSWER EVALUATION”** is submitted by **SHAD IBRAHIM N P (TKM20MCA-2034)** to the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the degree of Master of Computer Applications, is a bonafide record of the project work carried out by him under our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Internal Supervisor

Head of the Department

External Examiner

ACKNOWLEDGEMENT

First and foremost I thank GOD almighty and my parents for the success of this project. I owe sincere gratitude and heart full thanks to everyone who shared their precious time and knowledge for the successful completion of our project.

I am extremely grateful to **Dr.Fousia M Shamsudeen**, Head of the Department, for providing us with best facilities.

I would like to thank our coordinator, **Prof. Vaheetha Salam** and project guide **Prof. Natheera Beevi M**, Department of Computer Applications, who motivated me throughout the project.

I profusely thank all other faculty members in the department and all other members of TKM College of Engineering, for their guidance and inspirations throughout our course of study.

I owe my thanks to our friends and all others who have directly or indirectly helped us in the successful completion of this project.

SHAD IBRAHIM N P

ABSTRACT

Exams are a reliable and well-liked method of assessing the amount of information students have learnt. Exam evaluation is a difficult and time-consuming undertaking. Teachers will find this assignment handy since it requires less time and effort thanks to the automated response script assessment. Despite the fact that instructors may now give computerised objective and short answer tests, it may be difficult to evaluate descriptive responses. The computer must take into account the crucial elements that a human assessor takes into account while manually examining descriptive replies. This research attempt offered an intelligent assessment platform that takes into account the question type, relevant keywords, structural, conceptual, and linguistic factors when analysing a response in order to get around these difficulties. Concept graphs, fuzzy string matching, grammar checking, and other natural language processing similarity metrics were used to assess the proposed model. For instructors and students, the system also creates customised feedback and analysis reports to support focused learning.

Contents

1	INTRODUCTION	1
1.1	Problem Definition	2
1.2	Objective	2
2	LITERATURE SURVEY	3
2.1	Purpose of the Literature Review	3
2.2	Related Works	4
3	METHODOLOGY	7
3.1	Proposed system	9
3.2	System Architecture	10
3.3	Software Requirement and Specification	13
3.3.1	FlaskAPI	13
3.3.2	Python	14
3.3.3	PyCharm	14
3.3.4	HTML	15
3.3.5	CSS	16
3.3.6	JavaScript	16
3.3.7	Android Studio	16
3.3.8	SQLyog	17
4	RESULTS AND DISCUSSIONS	19

5 CONCLUSION	23
5.1 Advantages	24
5.2 Future Enhancement	24
REFERENCES	25
APPENDIX	27
A Screenshots	27

List of Figures

3.1	Block diagram of the proposed system	9
3.2	Fuzzy String Matching score	11
3.3	Language Score	11
3.4	Expected Concept Map	12
3.5	Student's Concept Map	12
4.1	Overview of evaluating description answer	20
4.2	Staff User Interface	21
4.3	Question preparation	21
4.4	Student's User interface	22
A.1	Admin Login	27
A.2	Admin Homepage	28
A.3	Student UI	29
A.4	Student subject selection	30
A.5	Student result display	31

Chapter 1

INTRODUCTION

In contrast to objective tests, where responses are judged only on the basis of right or incorrect, the assessment of descriptive answers also takes into account the test-cognitive taker's process, making descriptive answers crucial for evaluating a student's knowledge. Descriptive responses evaluate a student's conceptual understanding as well as their writing and language abilities. Descriptive exam evaluation is time-consuming and challenging since the teacher grading the answer needs to have a thorough comprehension of the subject matter. The replies may differ depending on the vocabulary the students use and the way they are written, so the teacher cannot assume that there is only one solution that is the only appropriate response to the question. Therefore, it is impossible to conduct a greater number of descriptive answer-based examinations due to the paucity of subject-matter experts. The majority of the time, string matching is used in computer programs to evaluate objective replies and short answers. However, it is important to include human intelligence for understanding the ideas and grade the replies while also taking into consideration linguistic elements like vocabulary, grammar, and syntax when analysing descriptive responses with a machine.

1.1 Problem Definition

The Descriptive Answer Evaluation project can be defined as the automated system for evaluating descriptive answers written by students. The system is useful in the education field as well as technological field. The project uses the Natural Language Processing technology to get the accurate mark for a particular question. The major motivation behind choosing the project was the drawbacks of the existing system. The system that we have now is a complete manual one. And which is followed from the very beginning. The major drawbacks of the existing system are mentioned below,

- It takes longer to know the outcomes because manual review is a time-consuming process.
- A significant quantity of human resources are used to evaluate the descriptive answers.
- A detailed response that received high marks may receive a low score from a different assessor.
- Even the same assessor may award a different grade for the same response at different occasions depending on how they are feeling.

1.2 Objective

The project's major purpose is to:

- presents a system that use deep learning and NLP to automatically analyse response scripts.
- When analysing descriptive replies, it is imperative to take into account important factors including the kind of the question, the desired response, important keywords, synonyms, grammar, and conceptual comprehension.
- Attempt to minimise the time and effort required for the manual evaluation

Chapter 2

LITERATURE SURVEY

A literature review is an overview of all the writings that have been done on a certain topic. The phrase can refer to a whole academic document or a section of a larger academic work, such as a book or an essay. It is a compilation of all previous study on a topic. Academic books, journal articles, and other resources relevant to a certain subject of study are examined in the literature review. This previous work should be mentioned, summarised, critically assessed, and explained in the review. The literature review gives credit to the work of other researchers. This shows the reader that your work is well thought out. When a previous work in the field of study is mentioned, it is assumed that the author has read, evaluated, and used that work in the current work. A literature review gives the reader a "landscape" of the field, so that he or she can see how things have changed. This landscape shows the reader that the author has, in fact, taken into account all or almost all of the important works done in the field before.

2.1 Purpose of the Literature Review

1. It makes it simple for readers to conduct research on a certain subject by choosing excellent articles or studies that are pertinent, significant, vital, and valid and compiling them into a single comprehensive report..
2. It forces researchers to sum up, assess, and contrast original research in that particular field, giving them a great place to start when conducting research in a new field

3. It prevents researchers from repeating earlier findings.
4. It can suggest topics to focus on or give hints as to where future research is going.
5. It emphasises the important findings.
6. It points out gaps, discrepancies, and inconsistencies in the literature.
7. It offers a helpful critique of the methods and strategies used by other researchers.

2.2 Related Works

S. P. Kar, J. K. Mandai et al. [2] provide a method for grading students in online courses that requires the teacher to enter the question, the student's response, as well as any necessary and minimal keywords. The presence of the keywords in the provided answer is directly checked. As a similarity metric for scoring, edit distance is used. Edit distance measures how few changes need to be made in the student response for it to be similar to the desired response, but it ignores other factors of similarity and synonyms.

The method of creating vectors from the keywords collected from the answer is covered by I. Das et al. [3]. The teacher must enter the expected keywords, the questions, and the responses. These keywords are compared to the keywords in the anticipated response to create the 0/1 vectors. The similarity is examined using the cosine similarity and term frequency approaches. Since only scores of 0 or 1 are taken into account for keyword matching, there is no provision for partial marking.

Ramamurthy M., Krishnamurthi, et al. [4] have described a system that largely concentrates on language-dependent elements. Answer vectors based on TF/IDF are created when the NLP pre-processing is finished, including noun-verb, noun-verb-adjective-adverb, noun-verb synsets, and noun-verb-adjective-adverb synsets. To determine marks, these vectors are examined using cosine similarity to the predicted answer vectors.

The student answer and predicted response are split into word vectors to calculate the similarity score, in accordance with D. V. Paul et al. [9]. The similarity matrix, which is comprised of a similarity matrix and columns I, j , and it, represents the similarity score between the j th vector of the predicted response and the i th vector of the student answer. Pairwise comparisons of the student response and anticipated answer vectors are used to fill the similarity matrix. The greatest similarity score for each student's answer vector is taken into account when averaging all of these values to determine the student's final grade. They employed word order similarity to determine how similar two vectors are to one another.

An approach based on semantic similarity was presented by Meena K., Lawrance R., et al. [10] to assess descriptive responses. Following clipping and stemming, the answer is subjected to latent semantic analysis (LSA), which creates a matrix with various words represented as rows and paragraphs as columns. Self-organizing maps are used to group vectors (SOM). The cosine similarity is then used to compute the degree of similarity between two vectors, and the result eventually decides the student's final grade.

Md. Motiur Rahman et al. [6] employ the weighted average of the ratings provided by the different similarity measures, including cosine, Jaccard, bigram, and synonym similarity. Keyword-based summarization is used to preprocess the replies before cutting and stemming. To assess how closely the student's response corresponds with the expected response, various similarity criteria are used. According to the criteria for calculating the weighted average, the weights for the similarity metrics are manually allocated.

For the autonomous assessment of response scripts, Neethu George, Sijimol PJ, Surekha Mariam Varghese, et al. [1] have suggested an approach that blends deep learning with NLP. The relevant text features from the answer scripts are extracted and transformed into glove vector indexes during the pre-processing step. The LSTM RNN is then given the vectors. The semantic rendition of the answer will be included in the embedding vector from the LSTM layer. The output layer, fully connected neural network layer, and dense layer will all forecast the one-hot score based on this number. Since there is just one kind of exam, it is challenging to train the model for each one.

Piyush Patil et al. [11] were the ones who initially proposed the notion of categorising the final marks. Teachers provide particular information for each question as well as keywords into the system. They used a Naive Bayes classifier which was based on three categories—keywords, grammar, and factors particular to the question—to determine the final score of the response.

S. Ben Salem et al. [8] created brief concept map tests for online courses to evaluate student understanding. Both the intended idea map and a list of candidate concept names must be provided by the teacher. In answer to brief, open-ended questions, the students must write responses in their own words. From the teacher's list of keywords, the students must select a handful and use those keywords to create an idea map. Using Wu and Palmer's similarity measure, which takes into account the semantic component of a concept map, the expected and student's maps are contrasted. Instead of automatically creating the concept map using the replies in natural language, they used the concept maps that the instructor and students had already generated as input.

When analysing the descriptive response, Nandini et al. [12] took the question type into account. Questions have been divided into three groups: factual, inductive, and analytical. The answers are assessed using cosine and Jaccard similarity based on this classification. The word order, similarity, degree of contradiction, and keywords are all part of the grading process. The innovative concept of question type consideration is offered in this study, although it is not specifically described how it influences the evaluation technique.

The authors of A. Kaur et al. [7] have contrasted three automated assessment methods: Text-Text, Graph-Graph, and Text-Cmap. They have determined that the TextCmap strategy is superior to other strategies. The study gap is that when contrasting the student and instructor responses, they did not take into account the synonyms of the concept map nodes.

Chapter 3

METHODOLOGY

The suggested system operates according to the Fig. 3.1 flow diagram. For the system to evaluate a student's descriptive answer, the teacher must provide the question, the intended response, and the predicted important keywords. The sort of question posed will determine how descriptive answers are evaluated [12]. The proposed framework divides issues into three groups: conceptual, particular, and analytical. For instance, your answer should be assessed more conceptually if the question asks you to "explain" or "speak" about a certain topic. The response should be brief if the inquiry is simple (explain, state), analytical (compare between, compare and contrast, or another comparable term), or both. The answer should, but need not, precisely match the one intended. Thus, the query falls within one of these categories. Extractive summary is used to compress answers that are excessively lengthy. Due to the database's storing of this preprocessed answer, the projected keywords, and the question type, preprocessing for the anticipated response is not required for every student's response. The same techniques are used to preprocess the responses from each student. The student's answer is then compared to the anticipated response using one of the four different methods outlined below.

1. Keyword Matching score: This strategy uses two methods, the first is to count the amount of anticipated keyword words that are included in the student's response, and the second is to calculate the cosine similarity score between the predicted and actual responses.
2. Language Score: The student's language proficiency is a significant factor in determining their final grade. Every time a learner makes a mistake, feedback and correction are given,

and the language score is adjusted as necessary.

3. Fuzzy String Matching Score: In natural language processing, fuzzy string matching evaluates whether two strings roughly rather than precisely match one another. It is utilised in this instance to determine how closely the student and teacher's responses line up. In other words, for responses that partially match, it will provide a rough matching score.
4. Concept Graph Score: Many authors [5][7][8] evaluate descriptive conceptual answers using the concept graph approach. Some writers used the teacher's input to create this idea graph, but the suggested system will create the concept graph on its own. Concept graphs of the predicted response and the student's response are created and compared to one another.

Because there is no one right response for a descriptive answer, the ontology learning module addresses the need of taking synonyms into account.

3.1 Proposed system

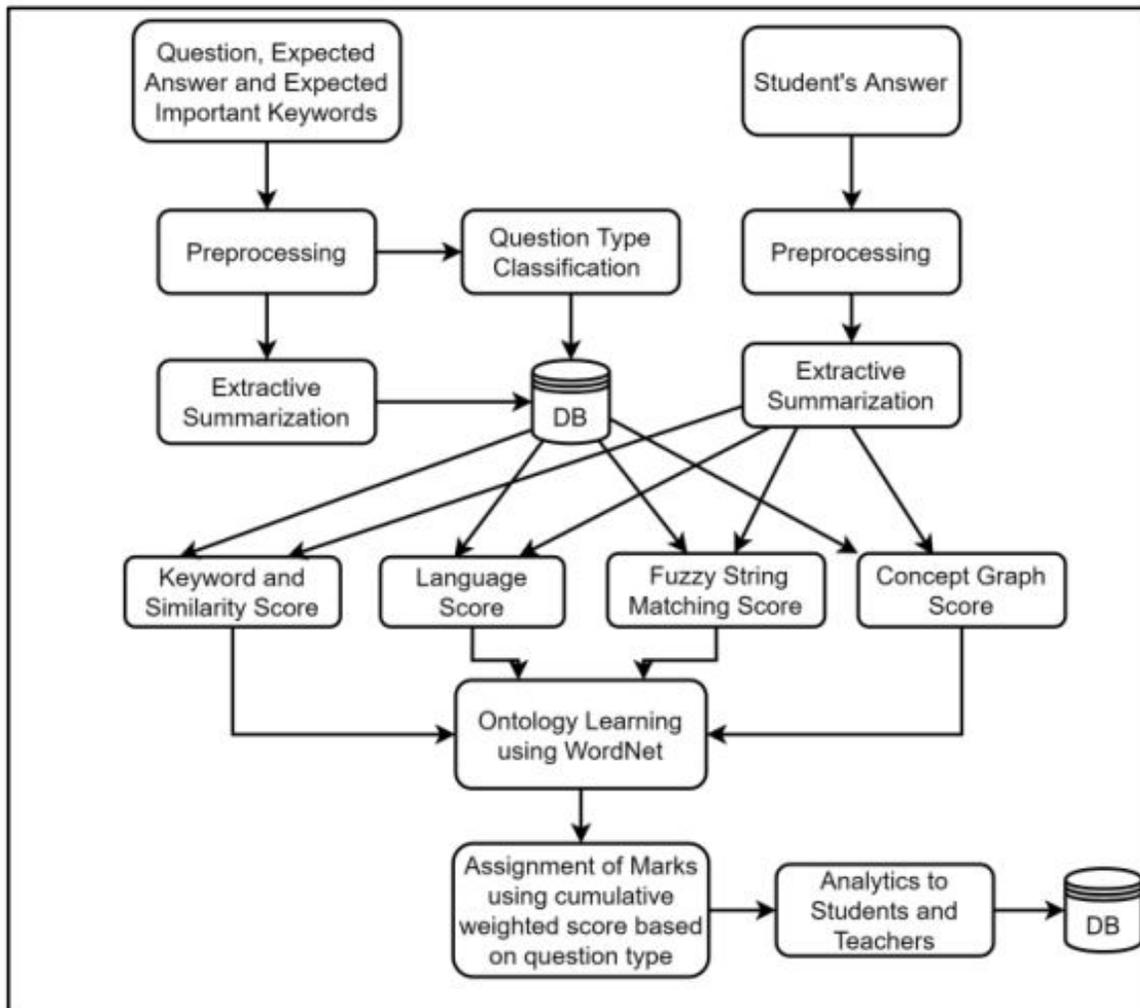


Figure 3.1: Block diagram of the proposed system

When a teacher manually evaluates a response, he or she first searches for the question type that is related to that specific response before analysing the response in accordance with the question type offered. The suggested technique takes into account this intelligence of first determining the question type and then choosing which factors should be given more weight for a certain question type. The weights assigned to each strategy are determined by the kind of inquiry. The weighted sums of the individual scores for each of the four different tactics are added to determine the final

score.

3.2 System Architecture

The system is web based. The input is taken from the teacher's side through a web page and the input is passed to the python program running in the server side. The server program perform tasks such as preprocessing and feature extraction using NLP on the input data of both student and examiner(teacher),and the cosine similarity of these keywords will be calculated and gives the appropriate mark to student.The student takes the examination through android application.

1. **Input Module** : It is possible for a teacher to design an exam, in which case they must provide each question, the desired response, a list of key phrases, and the marks they expect students to receive for that particular question. By giving the stated information, the teacher can directly design the exam using the suggested system.
2. **Preprocessing Module**: When natural language is involved, prepossessing is an important step that needs to be taken. Python's NLTK library is used to break the answer into sentences. If the answer is too long, it is summed up so that only the most important sentences can be taken out. The other sentences are just filler, so they can be ignored since doing more work takes a long time
3. **Keyword and Similarity Module** : For the suggested method, each student's response is compared to the significant keywords that a teacher entered for a particular answer in the database.
4. **Fuzzy String Matching**: Fuzzy string matching is a way to find strings that are similar to a given string but not exactly the same. Fuzzy string matching helps find the right word when a user misspells a word or only types part of it. The fuzzy string matching algorithm doesn't just look at whether or not two strings are the same. Instead, it measures how close two strings are to each other. Most of the time, this is done with a distance metric called "edit distance." This shows how close two strings are by figuring out how few changes need to be made to change one string into another.

```
Teacher Answer: Databases are of two types.  
Student Answer: There are two types of databases.  
Fuzzy String Matching Score Percentage: 100.0
```

Figure 3.2: Fuzzy String Matching score

5. **Language Module** :The language module of the proposed system tells the student what he or she did wrong with grammar or spelling and suggests better sentences, as shown in figure.

```
Mistakes found:  
Possible spelling mistake found.  
Offset 16, length 10, Rule ID: MORFOLOGIK_RULE_EN_US  
Message: Possible spelling mistake found.  
Suggestion: receives; recitatives  
The left atrium recieieves oxygen-rich blood from the lungs and pu...  
                ^^^^^^^^^^^  
Corrections: The left atrium receives oxygen-rich blood from the lungs and pumps it
```

Figure 3.3: Language Score

6. **Concept Graph Module**: Concept Graph is a directed graph made up of ideas, which are called nodes, and the connections between them, which are called edges. If there is a connection between nodes A and B, called an edge, then concept A is related to concept B through that connection. The first step in making a concept graph from textual data is to break the data into sentences. For each sentence, you should pull out the ideas and a connection between them. Usually, a sentence that stands on its own has a subject, an object, and a verb.

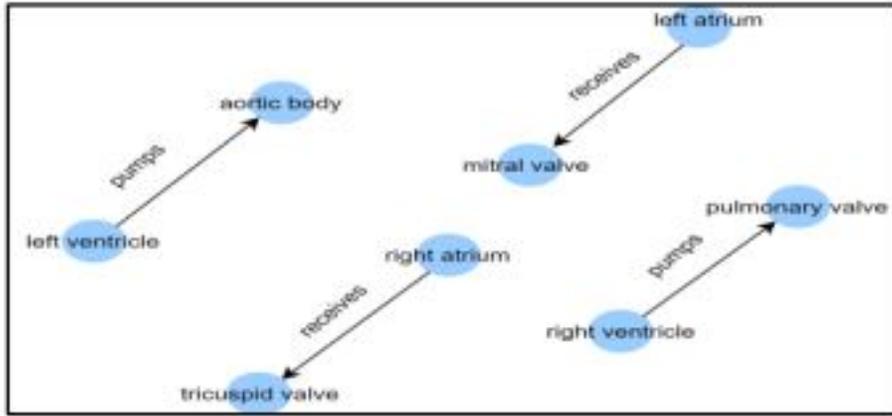


Figure 3.4: Expected Concept Map

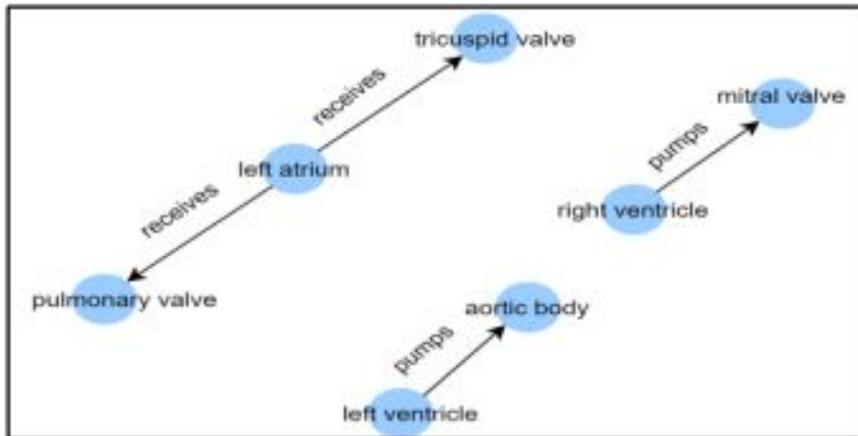


Figure 3.5: Student's Concept Map

3.3 Software Requirement and Specification

The tools used for the project are :

- Flask API
- Python
- PyCharm
- HTML
- CSS
- Javascript
- Android Studio
- SQLyog

3.3.1 FlaskAPI

Python-based Flask is a microweb framework. Due to the fact that it doesn't need any special tools or libraries, it is referred to as a microframework. It lacks any components that current third-party libraries already use to offer common capabilities, including a database abstraction layer, form validation, or other components. The Flask framework does, however, include extensions that may be used to add application functionality in the same way that they were integrated into the core framework. Extensions are provided for tools connected to common frameworks, such as object-relational mappers, form validation, upload processing, and numerous open authentication protocols. It was developed to allow Python programmers to create web applications. Its design prioritises speed, simplicity, and scalability for demanding applications (which is a lightweight application that provides a narrowed list of features with a well-defined contract). Flask, which Armin Ronacher created by encasing Werkzeug and Jinja, reminded me more of an April Fool's prank. They were surprised by the positive feedback, which motivated them to do better. Flask is adaptable and flexible. Unlike some other web frameworks, Flask's majority of components can be changed, but this is extremely unlikely for certain others. Because of its integrated support, built-in

development server, and other features, Flask offers unit testing and would enable you to transition to a web framework by changing some extensions. Because of its simplicity, Flask is particularly user-friendly for beginners, providing them room to grow as programmers. It also makes it possible for developers to construct apps quickly and easily.

3.3.2 Python

Python is an interpreter-based, high-level, general-purpose programming language. The architecture of Python prioritises code readability and makes extensive use of whitespace. It offers the fundamental building blocks needed for simple programming on both small and big scales. Python has an independent memory management system and a dynamic type system. It supports several programming paradigms, including imperative, functional, procedural, and object-oriented, and has a large and comprehensive standard library. Python is a multi-paradigm programming language. Many of its qualities enable functional programming, aspect-oriented programming, object-oriented programming, structured programming, met programming, and met objects (magic methods). Numerous other paradigms, such as design by contract and logic programming, also have extensions available. Python's memory management system consists of reference counting, a cycle-detecting garbage collector, and dynamic typing. Furthermore, it has the ability to dynamically resolve names (late binding), which binds variable and method names as the programme is being executed. Instead of using curly brackets or keywords to delineate blocks, Python uses whitespace indentation. The indentation is increased after some sentences and dropped at the end of the current block. As a consequence, the program's visual organisation accurately depicts its semantic structure. Additionally infrequently, this function. Off-side rule is how it's referred to.

3.3.3 PyCharm

Python's integrated development environment (IDE) is Pycharm, a cross-platform piece of software. It is well-liked since it supports the majority of Python versions and is compatible with the Linux, macOS, and Windows operating systems. It is a well-liked programme for

developing Python apps. The intelligent code editor in PyCharm makes it much easy to write good Python code. To make the code easier to understand and read, it employs unique colour schemes for keywords, classes, and functions in addition to syntax and error highlighting. In addition to having the ability for intelligent code completion, the code editor generates instructions for wrapping up the present code. It is much simpler to identify issues and errors, integrate linters, and develop quick fixes.

- PyCharm is a simple-to-use integrated development environment that requires little setup.
- PyCharm comes with a number of useful plugins and time-saving shortcuts. It combines IDE and library features, like as colorization and auto-complete.
- We may view the source code with only a single click.
- With PyCharm, software development is greatly accelerated..
- Being able to identify code issues improves the development process even further.
- The Python developer community is very broad, so we can answer any queries or concerns very away.

3.3.4 HTML

The most widely used language for constructing web pages is HTML, or Hyper Text Markup Language. Hypertext refers to the connecting of web pages (HTML documents). Links on a webpage are hence referred to as hypertext. You may use HTML to simply "mark up" a text document with tags that tell a web browser how to structure it for display since, as its name implies, HTML is a markup language. When HTML was first developed, it was done so to let academics communicate scientific data more easily and to design text structures like header paragraphs and lists. Web pages are now routinely prepared in HTML utilising a variety of tags.

3.3.5 CSS

The creation of visually beautiful web pages is made easier with CSS, commonly known as cascading style sheets. A web page's appearance and feel are managed through CSS. The text's colour, font style, distance between images or colours, layout designs, and display changes for various devices and screen sizes may all be controlled using CSS. Even though it's simple to grasp and use, CSS allows users extensive control over how an HTML text is rendered. The most common way that CSS is used as a markup language is in conjunction with HTML or XHTML.

3.3.6 JavaScript

The programming language JavaScript, sometimes referred to as JS, is one of the fundamental building blocks of the World Wide Web, along with HTML and CSS. For client-side web page behaviour, JavaScript is employed by over 97 percent of websites, and third-party libraries are often used. A specialised JavaScript engine is built into every major web browser to execute code on consumer gadgets. JavaScript is a high-level, often just-in-time, compiled programming language that complies with ECMAScript. It has first-class functions, prototype-based object orientation, and dynamic typing. Imperative, functional, and event-driven programming paradigms are all supported. Regular expressions, dates, text, the Document Object Model, and regular expressions are all supported via APIs (DOM). Now a common feature of many servers, JavaScript engines were formerly mostly available in web browsers.

3.3.7 Android Studio

Android Studio is the name of the authorised integrated development environment (IDE) for building Android apps. It incorporates the developer and code editing tools from the Java application integrated development environment IntelliJ IDEA.

Android Studio uses a Gradle-based build system, emulators, code templates, and Github integration to make it easier to create Android OS apps. Modalities are groups of source

code and resource files that are present in every project in Android Studio. Modules for Android apps, modules for Google App Engine, and library modules are some examples of these modalities.

The Instant Push feature of Android Studio is used to push updated code and resources to running applications. A code editor offers code completion, refraction, and analysis in addition to helping the developer write code. The APK file format is then used to submit applications created using Android Studio to the Google Play Store.

The software was initially introduced in May 2013 at Google I/O, and the first stable edition was made available in December 2014. Desktop PCs running Mac, Windows, and Linux can use Android Studio. It assumed the position of Eclipse Android Development Tools as the primary IDE for creating Android applications (ADT). Google offers the Software Development Kit and Android Studio for direct download.

3.3.8 SQLyog

The most effective management, admin, and GUI tool for MySQL is SQLyog Ultimate, which combines functionality from the MySQL Query Browser, Administrator, phpMyAdmin, and other Front Ends and GUI tools. SQLyog is a quick, easy, and compact graphical tool for controlling your MySQL databases. SQLyog Ultimate was developed for people who prefer MySQL as their RDBMS of choice. Regardless matter whether you like the control of handwritten SQL or want to work in a visual environment, SQLyog makes it simple for you to get started and provides tools to improve your MySQL experience. SQLyog Ultimate makes it simple to access frequently used SQL scripts. By using the SQLyog interface, the script files can be saved on a local disc or a shared network drive, or they can link to an existing file that is present everywhere Windows can access it. SQL scripts and file links may be organised using folders and subfolders.

With SQLyog, there are four primary connection techniques accessible. You shouldn't be hindered by any firewall, proxy, or ISP settings from administering your MySQL server properly (s). The following connection options are graded based on responsiveness and network efficiency.

- Direct connections utilising the MySQL C API are the easiest way to access and manage your MySQL server. Use this strategy instead if you don't worry about your MySQL traffic being faked. However, to do this, you must have direct access to your MySQL server. The majority of ISPs restrict or disallow remote access to the MySQL port.
- If you must encrypt the data in your MySQL traffic but still want the speed of the MySQL C API, use SSH tunnelling*. If your host allows SSH tunnelling, SQLyog can utilise the tunnel to connect to MySQL.
- HTTP Tunneling* may be able to help you if your ISP or MySQL provider bans remote access, either by blocking the MySQL port or by restricting access to "localhost." Since HTTP is stateless by design and the data is XML-encoded, the response is slower than with the other alternatives. However, you may still use all of SQLyog's wonderful features..
- Similar to HTTP Tunneling, HTTPS Tunneling creates a secure connection to offer an extra degree of security. When selecting this option, SSL Encryption, all communication between SQLyog and the MySQL server is encrypted. The MySQL server needs to support this option. Not all servers do.

Chapter 4

RESULTS AND DISCUSSIONS

The recommended technique considers the type of inquiry, the structural and conceptual elements of a response, significant expected keywords and their synonyms, and the linguistic features of an answer while performing an automatic evaluation. The factors mentioned above are used to determine each answer's final cumulative score. Teachers and students may use the database to obtain analytics since it contains the overall grade, the grade for each component, and the feedback. The automatic assessment procedure resembles the manual evaluation process more since it considers the question type, ideas, and synonyms. The recommended remedy makes it simpler for people to learn in a personalised fashion and addresses issues with the present manual evaluation approach.

As seen in Figure, a sample is created using the questions and anticipated responses for analysis. 4.1. The test to evaluate how effectively the recommended method operated included a wide range of question types. These comprised analytical, analytical, and conceptual inquiries. The columns with a grey background display the teacher's input, while the data in the other columns is generated by the proposed system. The final grade is calculated by multiplying the overall score ratio by the points allotted for the answers. The solutions in Fig. 4.1 receive a score out of five, but the total score is determined by the number of points allotted to the many questions. The teacher's input is shown in the columns with a grey background, while the data in the other columns is produced by the suggested system. The total score ratio is multiplied by the points allowed for the answers to determine the final grade.

The answers in Fig. 4.1 are graded out of five, but the overall grade is based on how many points are given to each of the several questions.

Question	Expected Keywords	Expected Answer	Student Answer	Question Type	Keyword and Similarity Score	Language Score	Fuzzy String Matching Score	Concept Graph Score	Final Marks(out of 5 mks)
How does blood flow through the heart?	left atrium, right atrium, right ventricle, tricuspid, pulmonary, aortic, mitral, valve, oxygen	The right atrium receives oxygen-poor blood from the body and pumps it to the right ventricle through the tricuspid valve. The right ventricle pumps the oxygen-poor blood to the lungs through the pulmonary valve. The left atrium receives oxygen-rich blood from the lungs and pumps it to the left ventricle through the mitral valve.	The left atrium is receives oxygen-poor blood from the body and pumps it to the right ventricle through the tricuspid valve. The right ventricle pumps oxygen-poor blood to lungs through the mitral valve. The left atrium receives oxygen-rich blood from the lungs and pumps it to the left ventricle through the pulmonary valve.	Conceptual	0.9311	0.75	1	0.25	1.903054

Figure 4.1: Overview of evaluating description answer

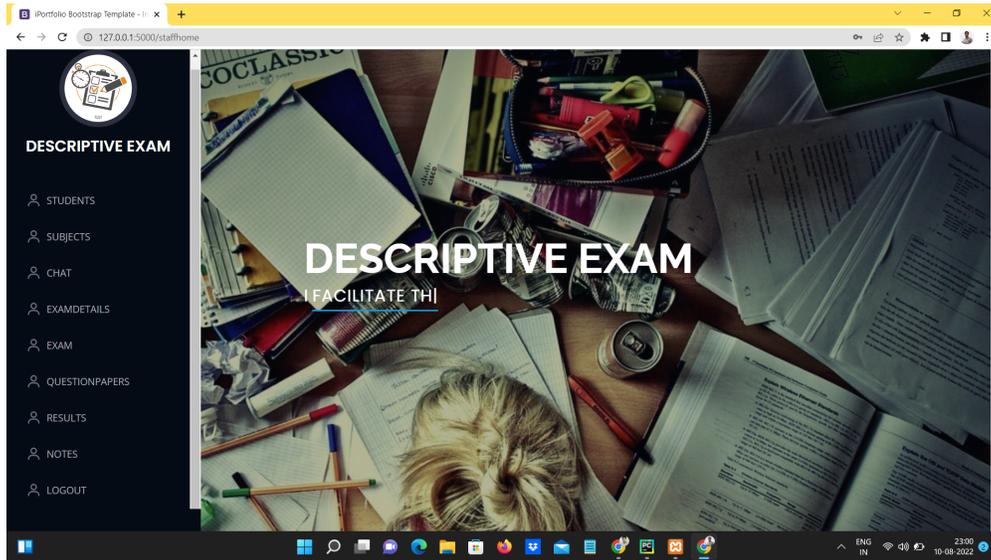


Figure 4.2: Staff User Interface

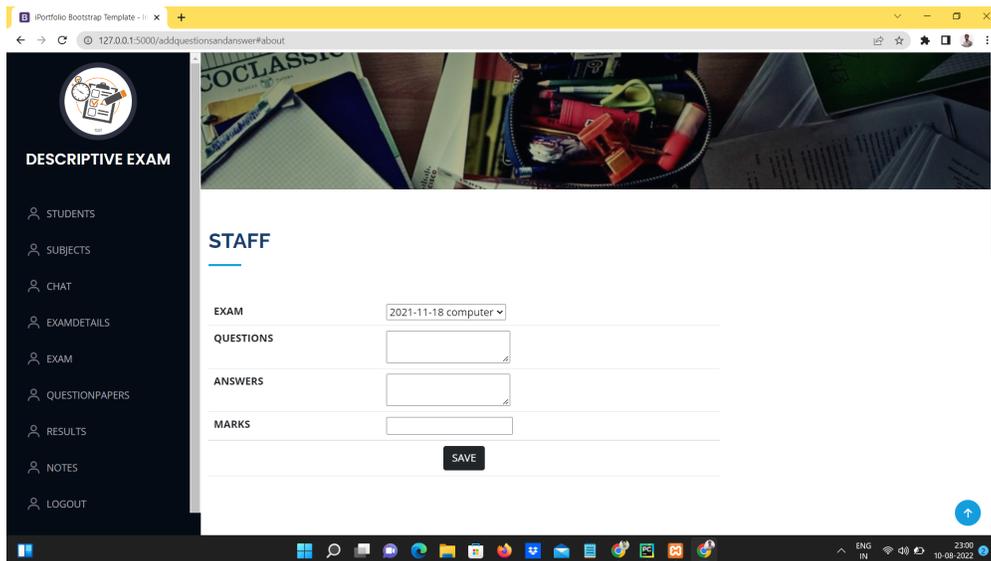


Figure 4.3: Question preparation

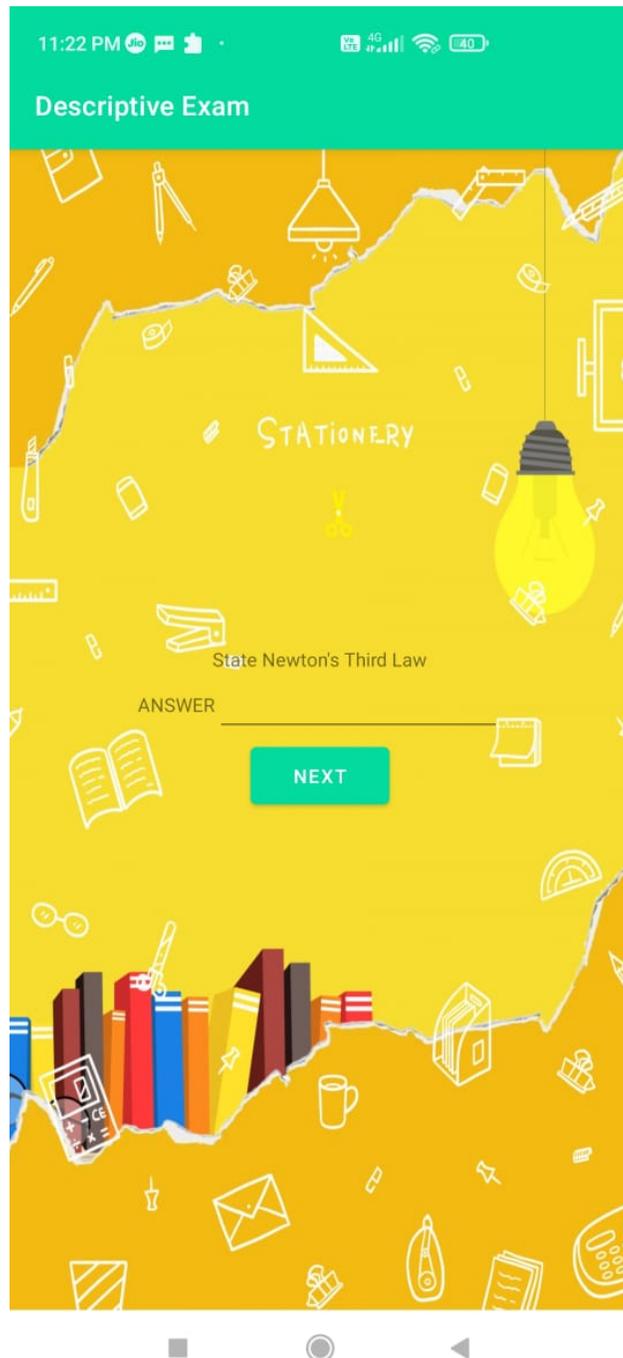


Figure 4.4: Student's User interface

Chapter 5

CONCLUSION

The Descriptive Answer Evaluation system is a great example of how Natural Language Processing can be used in the real world. It lets us automatically rate the answers to the descriptive questions while using fewer people. There are many ways to get it done. By comparing the answers given by the students to the ideal answer given by the examiner, who was probably a teacher, I tried to make an NLP model that could evaluate and predict the scores of the answers given by the students on a given topic. I used fuzzy string matching and the cosine similarity measure to find the similarities. The idea of automatically evaluating descriptive answers saves us the time and work that would be needed to review these answers by hand. The proposed system can be used for practise tests that are based on theory and only accept written answers. Answers including non-textual material, such as equations, diagrams, or tables, cannot be handled using the suggested approach. But in the future, by converting non-textual material into textual format, mathematical tests including formulae and illustrations might be scored automatically. In the future, a knowledge-based extraction mechanism might automatically draw keywords from a knowledge base and context that the instructor defines, replacing the proposed system's requirement that the teacher provide keywords. So that new types of questions can be added in the future, different score factors now have different weights based on which of three question categories they are in. Once this intelligent assessment platform has collected enough information, it can be used to predict students' interests in different academic fields and learn important things about how students

learn based on different learning outcomes.

5.1 Advantages

The main merits of proposed model are:

- Can save the time needed for the evaluation of descriptive answers. And this time can be spend more productive by the evaluators..
- The Human resource needed for the evaluation can be saved.
- The mood or mental situation of the evaluator no longer depend the score of the answer he/she evaluate.
- The pattern of online exams can be changed from objective to descriptive.

5.2 Future Enhancement

The feature scope of this particular system can be extended to multiple dimensions. Future development on this topic may lead to include more security aspects by using cryptographic techniques to enhance the integrity of exams. The system is more suitable for science subjects because of its objective nature. But in the future, the technique may be improved to be compatible with artistic languages like English and Hindi. Evaluation of the student's handwritten descriptive responses on a piece of paper is another potential improvement. It will be quite helpful in the educational field.

References

- [1] Neethu George, Sijimol PJ, Surekha Mariam Varghese, "Grading Descriptive Answer Scripts Using Deep Learning", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume.8 Issue-5 March,2019.
- [2] S. P. Kar, J. K. Mandai and R. Chatterjee, "A comprehension based intelligent assessment architecture," 2017 IEEE 6th International Conference on Teaching, Assessment and Learning for Engineering (TALE), HongKong, 2017, pp. 368-371, doi: 10.1109/TALE.2017.8252363
- [3] I. Das, B. Sharma, S. S. Rautaray and M. Pandey, "An Examination System Automation Using Natural Language Processing," 2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2019, pp.1064-1069, doi:10.1109/ICCES45898-2019.9002048.
- [4] Ramamurthy, Madhumitha Krishnamurthi, Ilango. (2016). Design and Development of a Framework for an Automatic Answer Evaluation System Based on Similarity Measures. Journal of Intelligent Systems. 26. 10.1515/jisys-2015-0031.
- [5] S. B. Salem, L. Cheniti-Belcadhi and R. Braham, "A concept map based scenario for assessment of short and open answer questions," 2017 International Conference on Engineering MIS (ICEMIS), Monastir, 2017, pp. 1-7, doi:10.1109/ICEMIS.2017.8273011.
- [6] Md. Motiur Rahman, Ferdusee Akter, "An Automated Approach for Answer Script Evaluation Using Natural Language Processing", IJCSET(www.ijcset.net) — 2019 — Volume 9, 39-47

- [7] A. Kaur and M. Sasikumar, "A comparative analysis of various approaches for automated assessment of descriptive answers," 2017 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, 2017, pp. 1-7, doi:10.1109/ICCIDS.2017.8272950.
- [8] S. Ben Salem, L. Cheniti Belcadhi and R. Braham, "Using concept maps in knowledge assessment of short and open answer questions," 2013 17th International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, 2013, pp. 646-650, doi: 10.1109/ICSTCC.2013-6679033.
- [9] D. V. Paul and J. D. Pawar, "Use of Syntactic Similarity Based Similarity Matrix for Evaluating Descriptive Answer," 2014 IEEE Sixth International Conference on Technology for Education, Clappana, 2014, pp. 253-256, doi: 15.1109/T4E.2014.60.
- [10] K. Meena and R. Lawrance, "Semantic similarity based assessment of descriptive type answers," 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), Kovilpatti, 2016, pp. 1-7, doi:10.1109/ICCTIDE.2016.7725366.
- [11] Piyush Patil, Sachin Patil, Vaibhav Miniyar, Amol Bandal, "Subjective Answer Evaluation Using Machine Learning", International Journal of Pure and Applied Mathematics — 2018 — Volume 118 No. 24
- [12] Nandini, V., Uma Maheswari, P. Automatic assessment of descriptive answers in online examination system using semantic relational features. J Supercomput 76,4430–4448 (2020). <https://doi.org/10.1007/s11227-018-2381-y>.

APPENDIX

A Screenshots

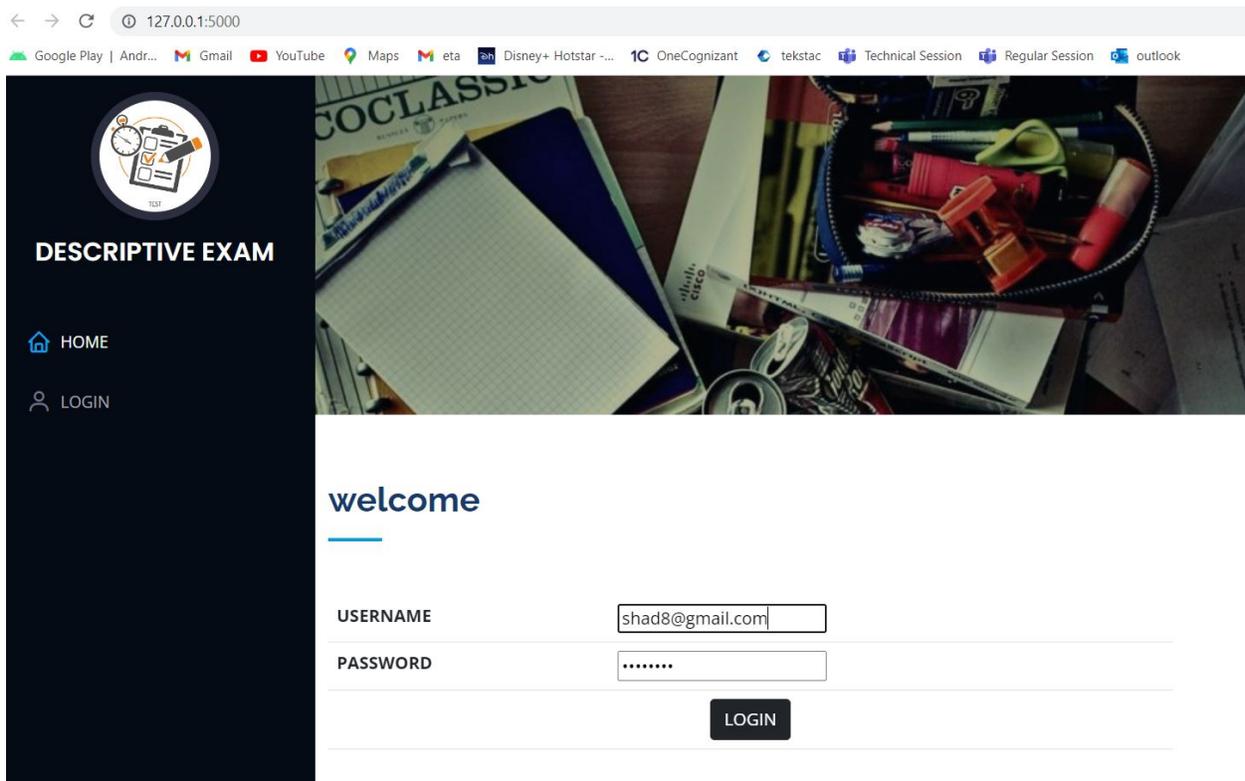


Figure A.1 : Admin Login

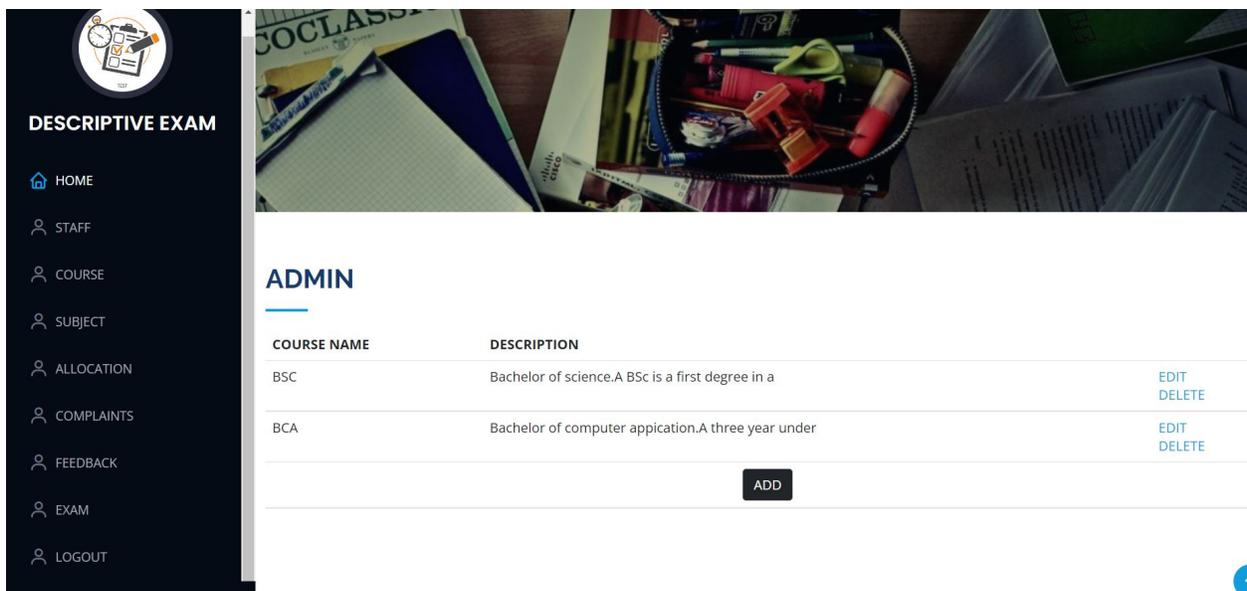


Figure A.2 : Admin Homepage

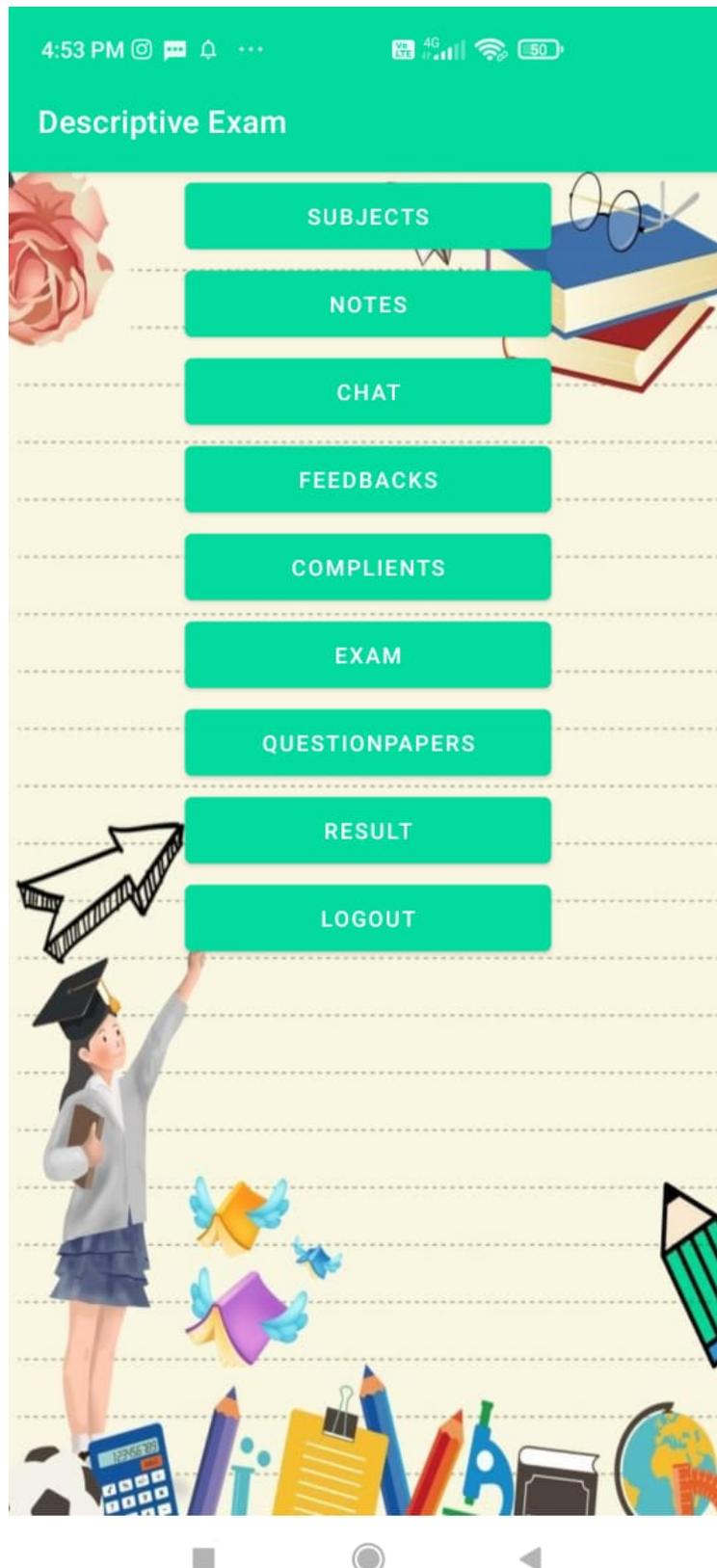


Figure A.3 : Student UI

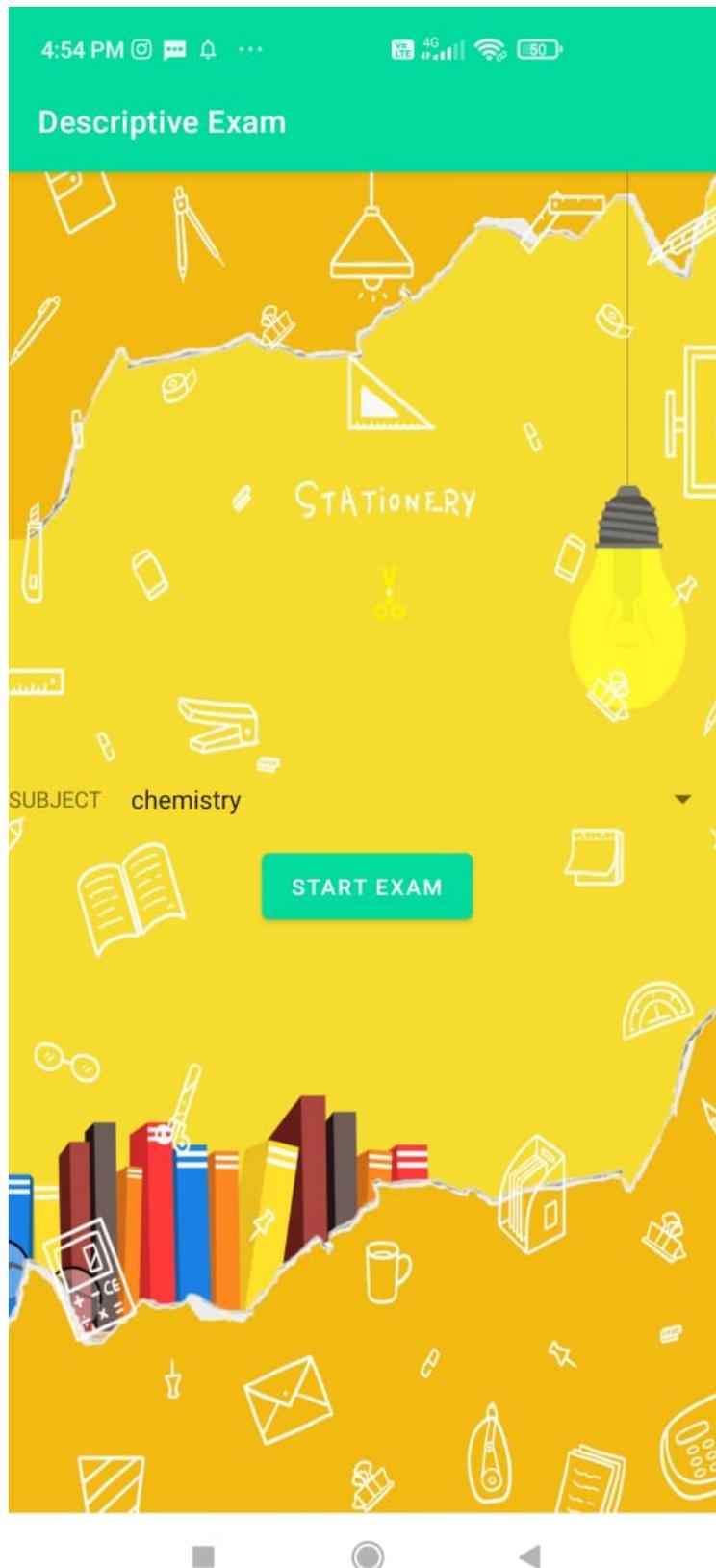


Figure A.4 : Student subject selection

