

Final Year Project Report
(Product Based)
Evaluate Hub



Project Advisor:
Ahtsham Ali

Submitted By:

Name	ID
Sawaira Iqbal	F2021332012
Aleena Sajid	F2021065333
Sameer Faisal	F2021065056

Session
2021 – 2025
BS SE

University of Management and Technology
C-II Johar Town Lahore Pakistan

Dedication

In the name of Allah, the Most Gracious, the Most Merciful.

All praise is for Allah Almighty, who blessed us with the strength, knowledge, and patience to complete this Final Year Project. We wholeheartedly dedicate this project to our beloved **parents**, whose prayers, sacrifices, and endless support have been our guiding light throughout this journey. We are deeply grateful to our **advisor, Sir Ahtsham Ali**, for his constant guidance, patience, and encouragement, which played a vital role in shaping our understanding and efforts. Our sincere thanks to the **University of Management and Technology (UMT)** and the **School of Systems and Technology (SST)** for providing us with the platform, resources, and environment needed to learn and grow. Lastly, we dedicate this achievement to all our teachers, mentors, and friends who stood by us during challenging times and helped us grow personally and professionally.

Final Approval

Panel of Examiners

- **Head of Department** _____
Department of Software Engineering
University of Management and Technology
Lahore
- **Program Director (Final Year Projects)** _____
Department of Software Engineering
University of Management and Technology
Lahore
- **Supervisor** _____
Department of Software Engineering
University of Management and Technology
Lahore
- **Co-Supervisor** _____

Acknowledgment

We would like to thank **Allah Almighty** from the bottom of our hearts for granting us the wisdom, strength, and perseverance to successfully complete this important milestone of our academic journey.

We are sincerely grateful to all the individuals who have walked this path with us. Most importantly, we thank our **professors and teachers**, whose continuous support and encouragement empowered us to build the skills and confidence needed to bring our project to life.

We also express our deep appreciation to the **School of Systems and Technology (SST)** and the **University of Management and Technology (UMT)** for providing us with the resources, mentorship, and a positive learning environment essential for our academic and professional growth.

A special word of thanks to our **project supervisor, Sir Ahtsham Ali**, whose guidance, patience, and insightful feedback were instrumental to the successful development and completion of our Final Year Project. His trust and belief in our abilities truly motivated us to push forward.

This work is dedicated to everyone who supported us throughout this journey. We sincerely hope that our efforts bring value, inspiration, and innovation for future generations.

Project Title

Evaluate Hub

Objective

To build an AI-based and automatic essay grading system with the capability of evaluating academic writing in a form of a widely used standardized test in higher education, such as IELTS, GRE or SAT. The system uses the power of the most advanced large-scale language models such as DeepSeek and LLaMA to effortlessly come up with correct band scores, valuable suggestions, and correction pointers; all according to the individual requirements of each testing platform. The platform will assist students, educators and institutions in improving the writing proficiency due to the intelligent, real-time analysis.

Undertaken by:

Name	ID
Sawaira Iqbal	F2021332012
Aleena Sajid	F2021065333
Sameer Faisal	F2021065056

Supervised by: Ahtsham Ali

Starting Date: October 2024

Completion Date: July 2025

Tools Used

- Programming Language: Python (TensorFlow,, Numpy, Pandas)
- Web: React.js / Node.js
- Model Train: Google Colab / Jupyter Notebook
- API Testing: Postman
- Coding Tool: VS Code
- Project Repository: GitHub
- Project Flow Management: Jira
- Frontend: HTML, CSS, JS
- Backend: NodeJS, Django

University of Management and Technology, Lahore

Similarity Report

Turnitin Originality Report

Evaluate Hub by Sameer Faisal, Sawaira Iqbal and Aleena Sajid

From Quick Submit (Quick Submit)

- Processed on 09-Jul-2025 19:07 PKT
- ID: 2712406342
- Word Count: 13112

Similarity Index

10%

Similarity by Source

Internet Sources:

8%

Publications:


2%

Student Papers:

9%

Sources:

1. 2% match (student papers from 30-Jul-2018)
Submitted to Higher Education Commission Pakistan on 2018-07-30
2. 1% match (student papers from 08-Mar-2021)
Submitted to Higher Education Commission Pakistan on 2021-03-08


Checked by


Verified by CLO

Note:

- Sometimes the overall similarity index may be a smaller than the repository percentages combined. This would be due to overlapping text within the repositories.
- It is a system generated report.

Declaration Form

I have carefully examined the documentation of the Final Year Project titled “*EvaluateHub*”; and I endorse that this documentation complies with the standards of an undergraduate level Final Year Project report.

The document has been checked for plagiarism through Turnitin software available in UMT Library. The similarities of the document are within acceptable range.

Moreover, the accompanying CDs contain PDF of the documentation, as well as the source code and binaries with user manual and installation guide.

FYP Advisor Name: _____

Signature: _____

Date: _____

Abstract

This project proposes **Evaluate Hub**, an AI-engineering essay evaluation system that is intended to turnover the evaluation of academic writing in international-recognized tests like **IELTS**, **GRE**, and **SAT**. This platform uses the potential of Natural Language Processing (NLP) and Generative AI to provide formatted band level scoring, feedback on mistakes and suggestions for improvement, and individual guides based on the content of the standard assessment rubrics.

In the process of development, a series of machine learning and deep learning models were tried out and compared - **Logistic Regression**, **Random Forest**, **AdaBoost**, and models based on the **BERT/LSTM** architecture. Although they all proved to give useful insights, their level of consistency in evaluation as well as contextualization was not adequate.

As a step of enhancing performance, we eventually introduced **DeepSeek** and **LLaMA** in the system with time replacing simple large language models. These models were very useful in improving the accuracy of scoring, detection of relevance and level of feedback hence its perfect contention in our final deployment.

Evaluate Hub is both scalable and intelligent that reduces manual grading to the minimum, assists learners in real time and provides consistent results. The system does not only allow an equal evaluation but it also promotes the student development with valuable mentorship. The next features will add support to speaking and listening modules, real-time analytics, and multicultural essay scoring.

REVISION CHART

Version	Primary Author(s)	Description of Version	Date Completed
<i>Draft</i>	All Members	Initial draft created with project title, objective, and basic structure	10 June 2025
<i>Preliminary</i>	All Members	Added Introduction, Domain Analysis, early model testing results (Logistic, RF, AdaBoost)	15 June 2025
<i>Final</i>	All Members	Full system workflow added with evaluation results from advanced models (BERT, LSTM)	20 June 2025
<i>Revision 1</i>	All Members	Integrated Llama model results and improvements to accuracy section	24 June 2025
<i>Revision 2</i>	All Members	Integrated Llama model, updated all comparisons, finalized result interpretation	27 June 2025
<i>Revision 3</i>	All Members	Final formatting, grammar check, and supervisor's feedback incorporated	28 July 2025
<i>Revision 3</i>	All Members	Final formatting, grammar check, and supervisor's feedback incorporated	31 July 2025

CONTENTS

CONTENTS.....	X
Definitions and Acronyms.....	XII
List of Figures.....	XIII
List of Tables.....	XIV
1. INTRODUCTION.....	1
1.1 Motivations.....	1
1.2 Project Overview.....	1
1.3 Problem Statement.....	2
1.4 Objectives.....	3
2. DOMAIN ANALYSIS.....	3
2.1 Customer.....	3
2.2 Stakeholders.....	4
2.3 Affected Groups with Social or Economic Impact.....	4
2.4 Dependencies / External Systems.....	5
2.5 Reference Documents.....	5
2.5.1 Related Projects.....	5
2.5.2 Feature Comparison.....	6
3. REQUIREMENTS ANALYSIS.....	7
3.1 Requirements.....	7
3.1.1 Functional Requirements.....	7
3.1.2 Non-Functional Requirements.....	8
3.1.3 Technological Requirements.....	8
3.2 List of Actors.....	9
3.3 List of Use Cases.....	9
3.4 System Use Case Diagram.....	10
3.5 Extended Use Cases.....	10
3.6 User Interfaces (Mock Screens).....	12
4. SYSTEM DESIGN.....	17
4.1 System Architecture Diagram.....	17
4.2 Class Diagram.....	18
4.3 Sequence Diagram.....	18
4.4 System Flow Diagram.....	19
4.5 Collaboration Diagram.....	19
4.6 Data Dictionary.....	19
5. IMPLEMENTATION DETAILS.....	20
5.1 Development Setup.....	20
5.2 Deployment Setup.....	21
5.3 Algorithms.....	21
5.4 Constraints.....	22
5.4.1 Assumptions.....	22
5.4.2 System Constraints.....	22
5.4.3 Restrictions.....	22
5.4.4 Limitations.....	22

6. TESTING.....	22
6.1 Extended Test Cases.....	22
6.2 Decision Table.....	45
6.2.1 Code Snippet.....	46
6.3 Traceability Matrix.....	49
6.3.1 RID vs UCID (requirements vs use cases)	49
6.3.2 Prototypes (RID vs PID)	50
6.3.3 Test Cases (RID vs TID)	50
6.3.4 Coverage (UCID vs TID)	52
7. RESULTS / OUTPUT / STATISTICS.....	52
7.1 % Completion.....	52
7.2 % Accuracy.....	52
7.3 % Correctness.....	53
8. CONCLUSION.....	54
9. FUTURE WORK.....	55
10. BIBLIOGRAPHY.....	56
10.1 Books.....	56
10.2 Journals.....	56
10.3 Articles.....	56
10.4 Research Papers.....	56
10.5 Other References.....	56
11. APPENDIX.....	57
11.1 Glossary of Terms.....	57
11.2 Pre-requisites.....	58

Definitions and Acronyms

Table: List of acronyms and definitions

Acronym	Definition
UMT	University of Management and Technology
POS	Point of Sale
NLP	Natural Language Processing
LLM	Large Language Model
ML	Machine Learning
DL	Deep Learning
API	Application Programming Interface
UC	Use Case
FYP	Final Year Project
GRE	Graduate Record Examination
SAT	Scholastic Assessment Test
IELTS	International English Language Testing System
UI	User Interface

List of Figures

Figure 3.1	Sample Use Case Diagram	10
Figure 3.2	User Account Registration Interface	12
Figure 3.4	User Authentication Screen	12
Figure 3.5	User Interface	12
Figure 3.6	Test Types	13
Figure 3.7	IELTS Essay Submission Interface	13
Figure 3.8	Evaluation Result with Feedback	14
Figure 3.9	IELTS Scoring Rubrics Display	14
Figure 3.10	SAT Essay Submission Interface	15
Figure 3.11	SAT Evaluation Result with Feedback	15
Figure 3.12	SAT Scoring Rubrics Display	16
Figure 3.13	SAT Essay Suggestions and Improvement	16
Figure 4.1	System Architecture	17
Figure 4.2	System Class Diagram	17
Figure 4.3	System Sequence Diagram	18
Figure 4.4	System Flow	18
Figure 4.5	System Collaboration Diagram	19
Figure 6.1	Sending GRE Essays to Backend for AI Scoring	46
Figure 6.2	Logic for Submitting and Evaluating IELTS Essay Using AI	47
Figure 6.3	Logic for Submitting and Evaluating SAT Essay Using AI	47
Figure 6.4	API Route Configuration in <code>urls.py</code>	48
Figure 6.5	Simplified <code>views.py</code> Function for LLaMA2-Based Essay Evaluation	48

List of tables

Table 2.1	List of Stakeholders	4
Table 2.2	Feature Comparison	6
Table 3.1	Functional Requirements	7
Table 3.2	Non-Functional Requirements	8
Table 3.3	Future Plan	9
Table 3.4	Extended Use Cases	10
Table 4.1	Data Dictionary	19
Table 4.2	Tools and Technologies	20
Table 6.1	Test Case 1	23
Table 6.2	Test Case 2	24
Table 6.3	Test Case 3	24
Table 6.4	Test Case 4	26
Table 6.5	Test Case 5	27
Table 6.6	Test Case 6	28
Table 6.7	Test Case 7	29
Table 6.8	Test Case 8	30
Table 6.9	Test Case 9	30
Table 6.10	Test Case 10	32
Table 6.11	Test Case 11	33
Table 6.12	Test Case 12	34
Table 6.13	Test Case 13	35
Table 6.14	Test Case 14	36
Table 6.15	Test Case 15	38
Table 6.16	Test Case 16	39
Table 6.17	Test Case 17	40
Table 6.18	Test Case 18	41
Table 6.19	Test Case 19	42
Table 6.20	Test Case 20	42
Table 6.21	Test Case 21	43
Table 6.22	Test Case 22	44
Table 6.23	Decision Tree	45
Table 6.24	RID vs UCID	49
Table 6.25	RID vs PID	50
Table 6.26	RID vs TID	50
Table 6.27	UCID vs TID	52
Table 11.1	Glossary of Terms	57

1. INTRODUCTION

The rapid advancement in artificial intelligence has transformed how educational assessments are conducted across the globe. Automated systems now offer an efficient alternative to traditional, manual evaluation methods.

1.1 Motivations

The increasing need of fairness, consistent and efficient scoring in academic and professional testing situations has brought into focus the disadvantage of manual essay evaluation. The conventional grading systems tend to be very tiresome, not reliable and so much dependent on the judgment of the human beings. As the AI and Natural Language Processing progress, it is now possible to automate these evaluations with a high degree of success and fairness.

Evaluate Hub was inspired after noting that most students find it hard to get timely and structured feedback on their writing, particularly on standardized exams being taken such as IELTS, GRE, or SAT. The institutions also have difficulty in scaling down on the human-based evaluation when there are a large number of students.

The gap was identified by our team, which attempted to fill it and develop an AI-based platform capable of grading essays on the fly, providing them with feedback based on test rubrics, and eventually assist learners and instructors. This is our initiative to create a fusion of not only academic but also advanced technology and the real needs of the world through which most of the learners around the world can be benefited.

1.2 Project Overview

This project addresses the growing demand for fast, unbiased essay evaluation using artificial intelligence. It offers a reliable system that aligns with international testing standards and enhances writing assessment through automation.

I. Overview Statement

Evaluate Hub is a web-based AI system designed to automatically evaluate academic essays for international standardized tests such as **IELTS**, **GRE**, and **SAT**. The system provides users with immediate scoring, grammar feedback, coherence analysis, and suggestions for improvement using advanced language models like **DeepSeek** and **Llama**. It aims to bridge the gap between manual evaluation limitations and the growing demand for intelligent, scalable assessment tools.

ii. Customer

While this project is not contracted by an external organization, the primary intended users include:

- Students preparing for IELTS, GRE, SAT, etc.
- Educational institutions and academies offering test prep courses
- Instructors seeking automated tools for assessment
- Independent learners who want to track their writing progress in real time

iii. Goals

- Test administration with automated essay scoring as well as global test rubrics
- Create band level band scoring with dimension level feedback (grammar, coherence, vocabulary, etc.)
- Facilitate various forms of testing and future scale-up
- Provide feedback in a neat, user friendly interface
- Ensure that it takes less time to evaluate and has higher consistency than a manual grading

iv. System Functions

- Essay submission and input interface
- Real-time evaluation using AI models
- Score calculation for key metrics (e.g., cohesion, grammar)
- Band prediction for specific tests (IELTS, GRE, SAT)
- Display of result summaries, and suggestions
- User authentication (optional: login/register)

v. System Attributes

- **Scalable:** Ability to cope with high volumes of essays with streamlined APIs and model pipelines
- **Accurate:** DeepSeek and LLaMA give high quality and test-aligned output
- **User Friendly:** It has a contemporary interface that makes it user friendly by the student and the educators
- **Modular:** Contingent to integrate other kinds of test formats or modules of foreign language in future

1.3 Problem Statement

Even within the contemporary educational environment, essay writing belongs among the most important elements of standardized international tests, including IELTS, GRE, and SAT, among others. Nevertheless, the process of evaluating these essays is quite cumbersome and it involves human beings to evaluate essay heavily. Many times manual grading is slow, biased and unreliable when there are too many students who were graded.

In addition, a significant amount of students does not receive timely and structured feedback in the course of preparation. Current automated systems are either inaccurate, only applicable in one form of test, or they do not give any tangible advice on how to make a change.

The need is emerging to have a secure, smart and scalable system of evaluations to automatically check the quality of the writing in real time as well as match the global exam rubrics. Lack of this kind of solution means that learners are unsure of their preparation and that institutions are unable to scale evaluation processes.

Thus, the necessity of a solution such as Evaluate Hub is rather obvious and the creation of such a solution with the help of AI and other modern models of language processing can rapidly and fairly help students to improve their writing and institutions to streamline their processes of assessment and evaluation.

1.4 Objectives

This project will mainly aim at designing a smart and flexible essay assessment system, called the Evaluate Hub, to automate the grading aspect of academic writing in different international standardized examinations, such as IELTS, GRE and SAT.

Specific Objectives:

i. Learning to Automate Essay Evaluation

To create a system that would test real-time grading of essays according to AI models coherent with standardized scoring rubric.

ii. Multi-test Support

In order to make the system fit the various forms of tests (IELTS, GRE, SAT) with scoring feedback that is test specific.

iii. Connect the Modern Language Models

In order to benchmark and compare different ML/DL models (Logistic Regression, Random Forest, BERT, LSTM) and complete LLaMA and DeepSeek via performance.

iv. Use Structured Feedback

In order to obtain personalized, dimension by dimension feedback regarding grammar, coherence, vocabulary and task relevancy.

v. Creating a User Friendly Interface

In order to build an efficient, smooth front-end, in which the users will be writing essays and getting detailed results with visualized references.

2. Domain Analysis

Domain analysis of an essay evaluation system focuses on understanding exam requirements (IELTS, GRE, SAT) and mapping scoring rubrics to automated assessment criteria. It identifies user needs, evaluation parameters, and system constraints for accurate and fair essay grading.

2.1 Costumers

Potential Customers / Target Users:

- **IELTS, GRE, and SAT Test Takers**
Students preparing for international exams who need immediate and accurate feedback to improve their writing skills.
- **Coaching Centers and Language Institutes**
Organizations that offer training for competitive exams and want to automate and enhance their assessment processes.
- **University Writing Labs and Academic Support Centers**
Institutions looking to provide AI-powered writing assistance for enrolled students.
- **Independent Learners and Professionals**
Individuals aiming to improve academic writing for admission or employment purposes in English-speaking countries.

Evaluate Hub addresses the shared need across these user groups for **real-time, fair, and standards-aligned essay evaluation**, offering a solution that is accessible, scalable, and intelligent.

2.2 Stakeholders

Stakeholders include students, teachers, and institutions who rely on accurate essay evaluation for learning and admissions. They influence system requirements and benefit from reliable, unbiased feedback.

Table 2.1: List of Stakeholders

Stakeholder	Role in System
Students	Primary users of the system. They input essays, receive evaluation scores and feedback.
Test Preparation Institutes	Use the system to automate essay evaluation and monitor student progress.
Instructors / Tutors	Use the system for guided writing practice, feedback analysis, and student improvement tracking.
FYP Team Members	System developers responsible for design, implementation, model integration, and testing.
Supervisor (Sir Ahtsham Ali)	Provides mentorship, monitors progress, and ensures academic alignment of the project.
University (UMT / SST)	Facilitates the project environment, technical support, and provides evaluation frameworks.
Future Developers / Maintainers	Responsible for scaling, updating models, adding test types, and improving accuracy.

2.3 Affected Groups with social or economic impact

All off these users and organizations will be affected socially and financially by the implementation of Evaluate Hub. The overview of the most affected groups is given below:

i. Students:

- Impact:** They can get access to free or cheap quality feedbacks that make them improve in their academic writing without going to costly tutors.
- Social Benefit:** Promotes selflearning, confidence enhancement as well as preparation to competitive examinations in a better way.

ii. Test Preparations Institutes:

- Difference:** Can decrease the reliance of manual marking enabling instructors to devote more time on individual instruction and less on script-like marking.
- Economic Advantage:** Cost savings of operations and better efficiency in the provision of quality education.

iii. Teachers / Instructors:

- Impact:** Automated feedback serves as an initial guideline to give additional instructions, saving time and guaranteeing equal estimates among students.
- Social Benefit:** Makes student evaluation more fair and consistence.

iv. Learners / Professionals:

- g. **Effect:** The professionals who are taking the route of international education or immigration are now in a position to gain access to instant feedback tools without getting involved in full-time coaching programs.

v. Academic Institutions:

- h. **Significance:** This system can be applied in writing labs or English classes where writing grading has to be normalized and there is a lack of student support systems.

2.4 Dependencies/ External Systems

Evaluate Hub relies on the LLaMA language model, Python-based NLP libraries, and a Node.js–Django architecture for real-time essay evaluation and feedback delivery.

vi. Large Language Models

- **DeepSeek API** – Used for essay evaluation, band scoring, and contextual feedback generation.
- **LLaMA Model** – Fine-tuned locally for performance benchmarking and comparison with other models.

vii. Development Tools & Libraries

- **Python (Scikit-learn, TensorFlow, Transformers)** – For training/testing traditional ML and deep learning models.
- **Pandas, NumPy, NLTK** – For data preprocessing and natural language feature extraction.
- **Matplotlib, Seaborn** – For visualizing performance metrics and result distributions.

viii. Frontend & Deployment

- **Streamlit / Node.js / React.js** – Web interface for user input, essay submission, and real-time feedback display.
- **Postman** – API testing and integration validation.
- **GitHub** – Version control and collaborative development.
- **Render / Localhost** – Deployment environment for model hosting and testing.

ix. Other External Systems

- **Google Colab** – Used for initial model prototyping and comparative evaluation.
- **Turnitin or Similar Tools** – (Optional) For future integration with plagiarism detection modules.

2.5 Reference Documents

During the development of **Evaluate Hub**, several research papers and previous works were consulted to better understand essay scoring frameworks, NLP techniques, and evaluation models. Below are selected related reference

2.5.1 Related Projects

x. Automated Essay Scoring using BERT and GPT Architectures

Authors: Kumar, A., & Sharma, P.

Summary: This paper discusses the implementation of transformer-based models for evaluating essays using coherence, grammar, and vocabulary measures.

Link: <https://arxiv.org/abs/2004.05212>

xi. Deep Learning Approaches to Predicting Band Scores in IELTS Writing Tasks

Authors: Lee, M., & Cho, J.

Summary: Focuses on using CNNs and RNNs to predict IELTS band scores. Also discusses rubric alignment and training data labeling.

Link: <https://aclanthology.org/2020.bea-1.3>

xii. A Comparative Analysis of Transformer Models for Text Scoring Tasks

Authors: Rahman, M., & Li, H.

Summary: Evaluates multiple transformer models (GPT, LLaMA, and T5) for effectiveness in automated scoring tasks. Supports choosing LLaMA as a top performer.

Link: <https://arxiv.org/abs/2305.09819>

2.5.2 Feature Comparison

Feature comparison highlights how the system surpasses traditional manual checking and existing tools by offering faster, unbiased, and rubric-based essay evaluation.

Table 2.2: Comparing Features

Sr No.	Comparison Feature	Existing Systems / Research	Evaluate Hub	Remarks
1	Multi-Test Support	The majority of the available systems are specific in one test format (i.e. IELTS only).	Supports IELTS, GRE, SAT, and can be extended to others.	More complex and reusable in varied academic situations.
2	Real-Time Feedback	There is some batch system where systems take feedback or evaluation at a later time.	Offers instant feedback and scoring on submission of an essay.	Is time saving and results in better engagement of students.
3	Model Accuracy and Depth	The normal models (such as Logistic Regression, Random Forest) can be applied.	Utilizes state-of-the-art LLMs (LLaMA, DeepSeek) to score more accurately.	Delivers higher accuracy and context-aware feedback.
4	Rubric Alignment	Limited alignment with standardized rubrics.	Feedback structured around official IELTS/GRE/SAT criteria.	Provides equitable, test focused assessment.

Sr No.	Comparison on Feature	Existing Systems / Research	Evaluate Hub	Remarks
5	Visual Analytics	The majority of the current solutions do not provide the evaluation results visualization.	Includes user-friendly breakdowns in scores.	Makes information more clear and easier to use.
6	User Interface	Most tools are API-only or not designed in an extremely user-friendly way.	Offers a web-based, interactive UI using Streamlit/Node.js/React.js /Django.	To be used both by institutions and individuals.

3. Requirements analysis

Evaluate Hub is designed to provide real-time, scalable, and rubric-based essay evaluation with a responsive interface and support for multiple test types.

3.1 Requirements

The core objective of this project is to build a smart and reliable system that evaluates essay-based answers in real-time using modern language models. The requirements outlined below were identified through a detailed analysis of user needs, academic testing standards, and the capabilities of AI-driven scoring systems.

This section presents the system’s functional, non-functional, and technical requirements, along with future development goals and interface expectations. These requirements guide the system’s structure and performance targets throughout the development process.

3.1.1 Functional Requirements

These requirements describe what the system should be able to do from a user’s perspective. Each one aligns with the purpose of delivering an automated and meaningful writing evaluation experience.

Table 3.1: Functional Requirements

RI D	Requirement	Type	System Behavior / Notes
FR 1	The system must evaluate essays by analyzing grammar, vocabulary, coherence, and structure.	Functional	Assess essays according to the chosen test's evaluation criteria (e.g., IELTS, GRE, SAT).
FR 2	The system should highlight specific areas where students need improvement.	Functional	Tailored suggestions must be generated based on individual performance.
FR 3	Teachers and mentors should be able to view student performance trends.	Functional	Analytics dashboard can show growth and weaknesses over time.
FR 4	Students must receive instant feedback after submission.	Functional	Feedback should be displayed within a few seconds of submitting the essay.
FR 5	The system must generate visual summaries (charts or graphs) of performance.	Functional	Graphs should break down scores across different writing dimensions.
FR	The evaluation must be consistent and	Functional	Models should be trained and fine-tuned

6	aligned with international grading standards.		according to official test rubrics.
FR 7	Users should be able to download their evaluation reports in PDF format.	Functional	Reports must include scores, suggestions, and visual analysis.
FR 8	The interface must allow simple and intuitive online essay submission.	Functional	Clear text input and test selection must be available to all users.

3.1.2 Non-Functional Requirements

These define the quality and efficiency standards the system must meet to function effectively in a real-world setting

Table 3.2: Non Functional Requirements

RI D	Requirement	Type	System Behavior / Notes
NF R1	The platform should support multiple essay evaluations at the same time.	Non-Functional	System must handle at least 50 essays/hour without lag or crashes.
NF R2	Essay evaluation should be processed and displayed within 10 seconds.	Non-Functional	Response time is critical to preserve real-time interaction.
NF R3	The interface must work across different devices and screen sizes.	Non-Functional	Fully responsive design for mobile, tablet, and desktop users.
NF R4	User authentication (optional) should be secure and session-based.	Non-Functional	Login/signup may be used to track user progress securely.

3.1.3 Technological Requirements

The technology requirements include a web-based platform with AI/ML models, and cloud or local hosting for scalable essay evaluation.

- **Languages and Frameworks:** Python (with Scikit-learn, TensorFlow, Transformers), Node.js/Streamlit
- **Libraries:** Pandas, NumPy, Matplotlib, NLTK
- **Tools Used:** Google Colab, Postman, GitHub
- **Optional Integrations:** PDF report generation, Firebase (if login is implemented)

Design Constraints

- The system must remain lightweight and deployable on both cloud and local machines.
- Essay scoring must be **transparent** and explainable through clear feedback.
- Each test (IELTS, GRE, SAT) must follow its official evaluation structure.

Interface Requirements

- Simple, clean user interface for essay input
- Dropdown menu to select test type
- Downloadable results with charts and improvement tips
- Optional login system to save and compare submissions

Table 3.3: Future Plan

Feature	Planned Enhancement
Listening Evaluation	AI model to assess understanding from user's written/audio response to a listening task
Speaking Evaluation	Scoring based on pronunciation, fluency, grammar, and coherence in speech
Reading Evaluation	Auto-marking of multiple-choice questions via answer-key logic

3.2 List of Actors

The following actors define the system's boundaries and represent all key users or systems that will interact with **Evaluate Hub**:

- i. **Student**
This is the primary user who writes and submits essays to the system, selects the test type, and receives feedback and scoring results. The student can also download evaluation reports.
- ii. **Instructor / Tutor**
This actor accesses feedback reports for multiple students and uses insights to analyze trends and guide their academic instruction. They may also monitor improvement over time.
- iii. **Admin / Developer**
Responsible for managing backend operations, model testing, bug fixing, and pushing updates. This actor may also access evaluation logs or raw model outputs for debugging purposes.
- iv. **AI Model (LLaMA / DeepSeek)**
While not a human actor, the model plays an internal system role. It processes essays, evaluates them, and generates scoring feedback based on the selected test rubric.
- v. **System Interface (Frontend Web UI)**
Serves as the bridge between users and models. Handles form submissions, displays feedback, and formats data visualizations. It also validates inputs before sending them for evaluation.

3.3 List of use cases

These are the primary actions that users can perform within the **Evaluate Hub** system: Submit Essay
Allows a student to type or paste their essay into the interface and send it for evaluation.

- **Select Test Type**
Enables the student to choose between IELTS, GRE, or SAT before essay submission.
- **Receive Evaluation Feedback**
The system shows grammar, coherence, and vocabulary scores along with band predictions and improvement suggestions.
- **Access Performance Trends**
Tutors or students can view graphs and summaries showing writing progress over time (if login is implemented).

- **Debug Evaluation Output**
Admin or developer can view raw model predictions for testing or performance validation.
- **Log In / Register (Optional)**
Allows users to create accounts, log in, and track their submission history.

3.4 System use case diagram

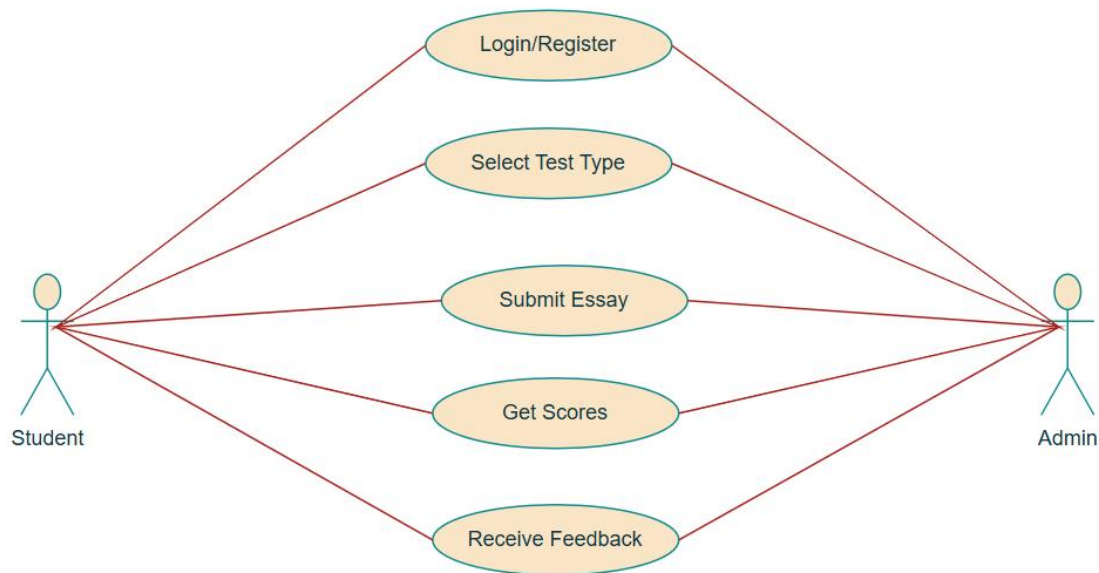


Figure 3.1 Sample use case diagram

3.5 Extended use cases

Table 3.4: Extended Use Cases

Use Case ID:	Enter a unique numeric identifier for the Use Case. e.g. UC-1.2.1		
Use Case Name:	Enter a short name for the Use Case using an active verb phrase. e.g. Withdraw Cash		
Created By:		Last Updated By:	
Date Created:		Last Revision Date:	
Actors:	Primary Actor: User (Student, Instructor) Secondary Actor: Evaluate Hub System (DeepSeek, LLaMA)		
Description:	The user submits an essay for automated evaluation. The system uses advanced language models to score and provide feedback based on standardized rubrics for IELTS, GRE, and SAT.		
Trigger:	User submits an essay via the chatbot interface.		
Preconditions:	1. System is accessible via web 2. User selects a test type 3. Essay text is provided 4. Models (LLaMA/DeepSeek) are functional		
Post conditions:	1. Essay is scored and feedback displayed 2. User receives dimension-wise scores and improvement suggestions 3. Report is downloadable		

Normal Flow:	<ol style="list-style-type: none"> 1. User accesses the Evaluate Hub chatbot. 2. User selects the test type (IELTS, GRE, SAT). 3. User types or pastes their essay. 4. User clicks “Submit”. 5. System validates input and test type. 6. Chatbot forwards essay to LLaMA/DeepSeek. 7. Model processes essay and generates scores. 8. Feedback (Grammar, Vocabulary, Coherence) is shown. 9. User may download the feedback report.
Alternative Flows: [Alternative Flow 1 – Not in Network]	<p>4a. Essay Input Missing In step 4 of the normal flow, if the user does not input an essay:</p> <ol style="list-style-type: none"> 1. The system shows: “Essay input is required.” 2. The user enters essay content. 3. The user clicks submit again and proceeds to validation and evaluation. <p>5a. AI Model Unavailable In step 5 of the normal flow, if the system cannot access the model:</p> <ol style="list-style-type: none"> 1. The system shows: “Evaluation temporarily unavailable.” 2. The user waits and submits the essay again when services are restored.
Exceptions:	<p>6a. Invalid Essay Content</p> <ol style="list-style-type: none"> 1. If special characters or unsupported input is detected, system rejects the input. 2. The system displays: “Invalid format detected. Please revise your essay.” 3. The user edits the essay. 4. The user submits the revised version for processing. <p>6b. No Test Type Selected</p> <ol style="list-style-type: none"> 1. The system shows: “Please select a test type before submitting.” 2. The user selects a test type and resubmits.
Includes:	<ul style="list-style-type: none"> • Essay Evaluation Engine • Rubric Comparator • Feedback Generator • Band Score Calculator
Frequency of Use:	This use case is expected to be executed frequently, especially during exam preparation seasons approximately 100–150 times per day.
Special Requirements:	<ul style="list-style-type: none"> • Evaluation must complete within 10 seconds. • Output must strictly follow test-specific rubrics. • System should support both mobile and desktop interfaces.
Assumptions:	<ul style="list-style-type: none"> • User understands the test format and instructions. • Essay input is written in English. • System has stable internet access and connection to models.
Notes and Issues:	<p>Future addition of speech and listening evaluation is planned.</p> <p>Plagiarism detection may be integrated later.</p> <p>Maximum essay size needs to be defined and validated (e.g., 500 words).</p>

3.6 User interfaces (mock screens)

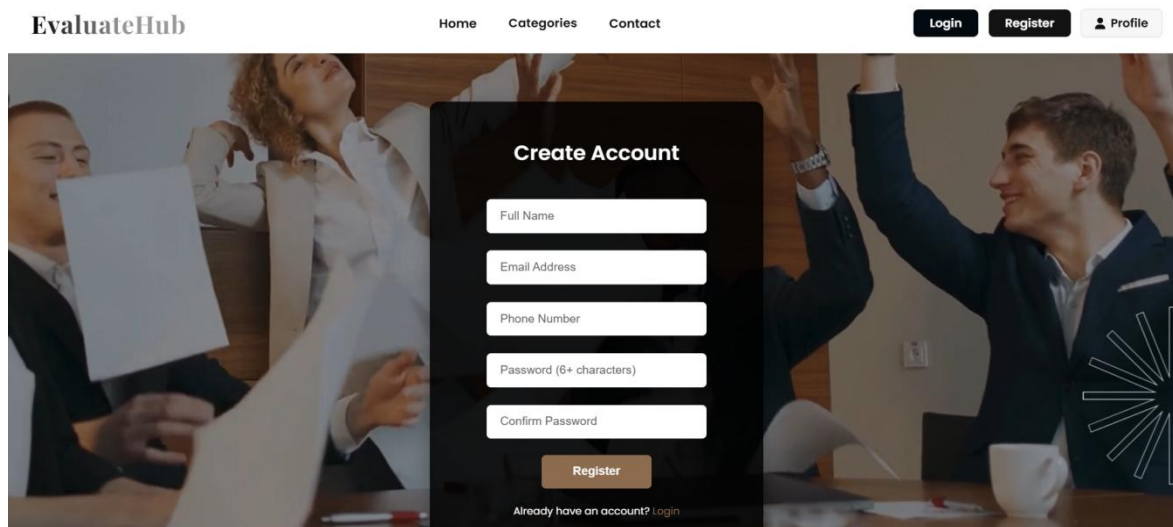


Figure 3.2 User Account Registration Interface

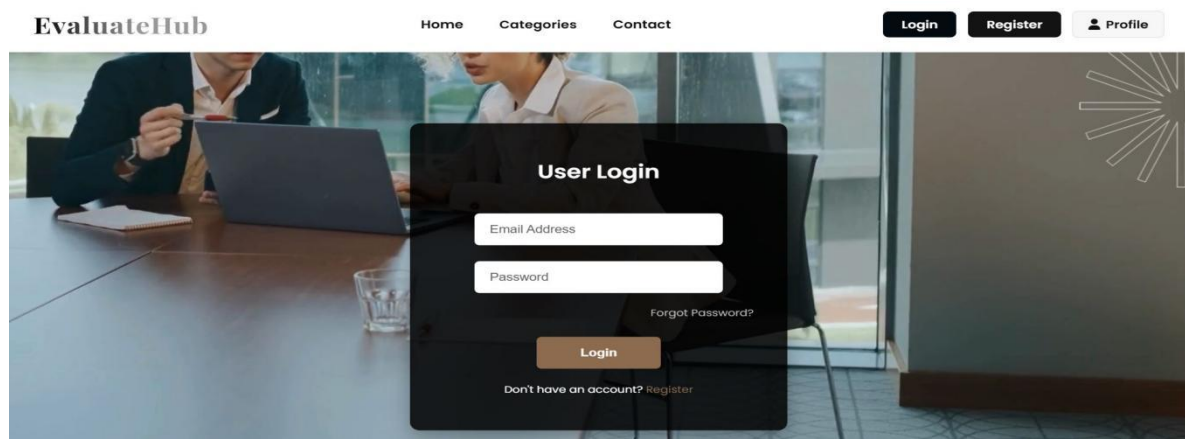


Figure 3.4 User Authentication Screen

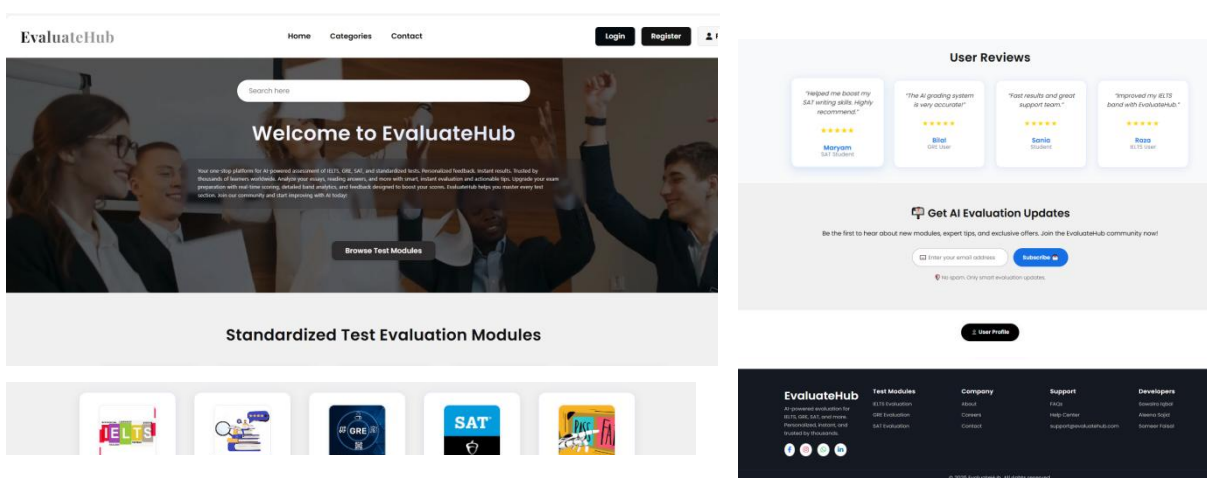


Figure 3.5 User Interface

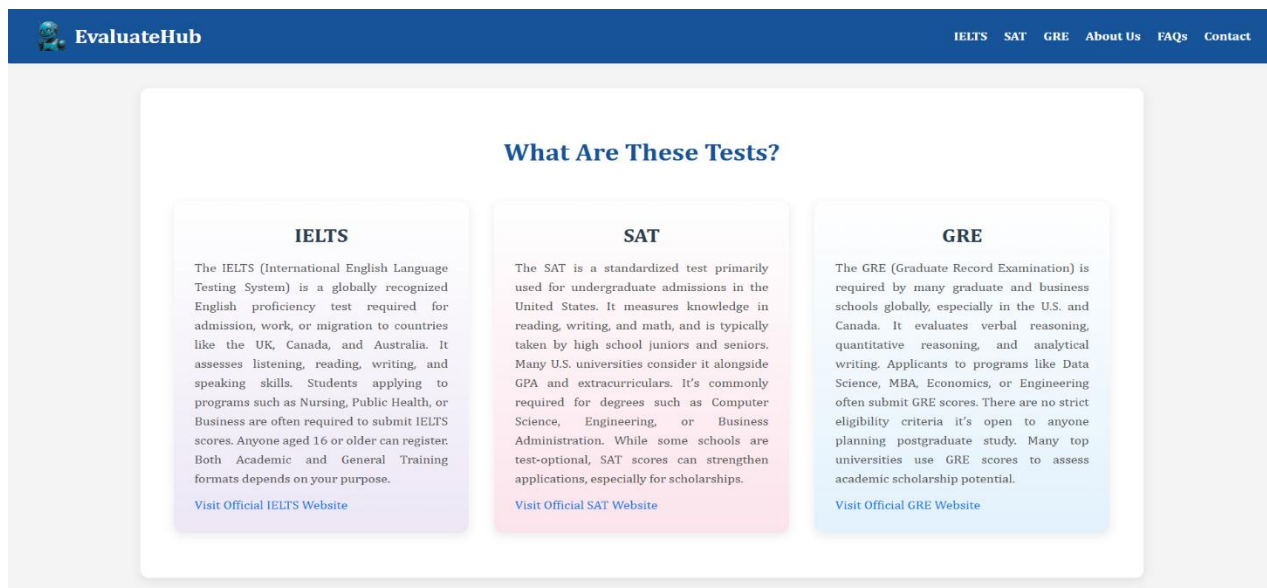


Fig 3.6 Test Types

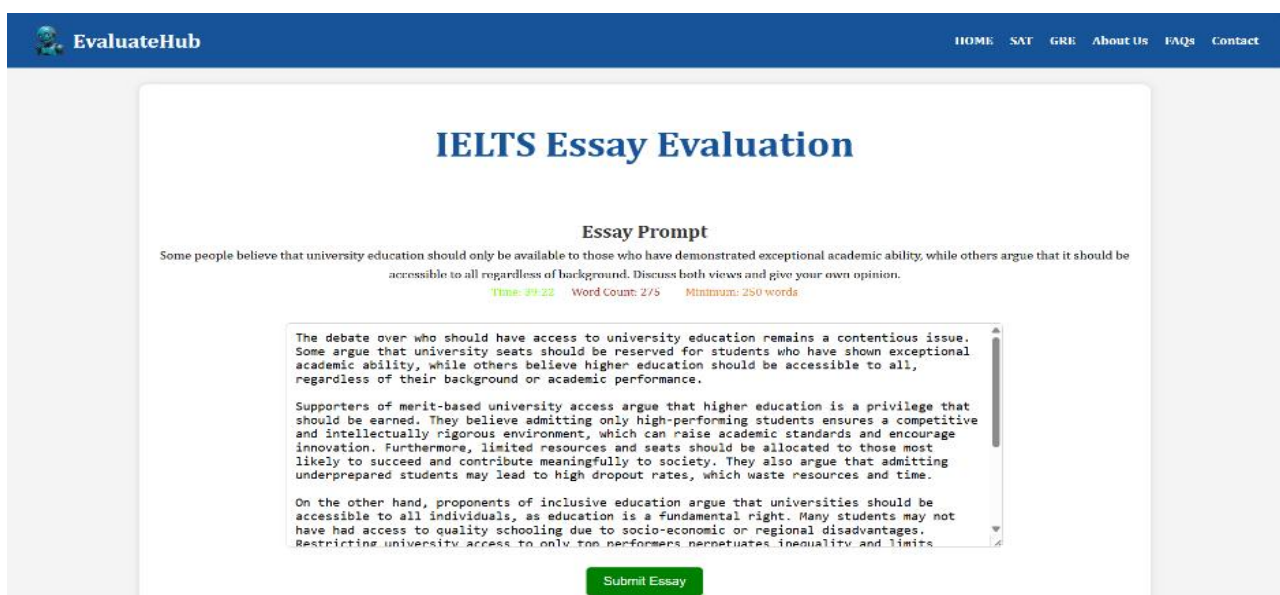


Fig 3.7 IELTS Essay Submission Interface

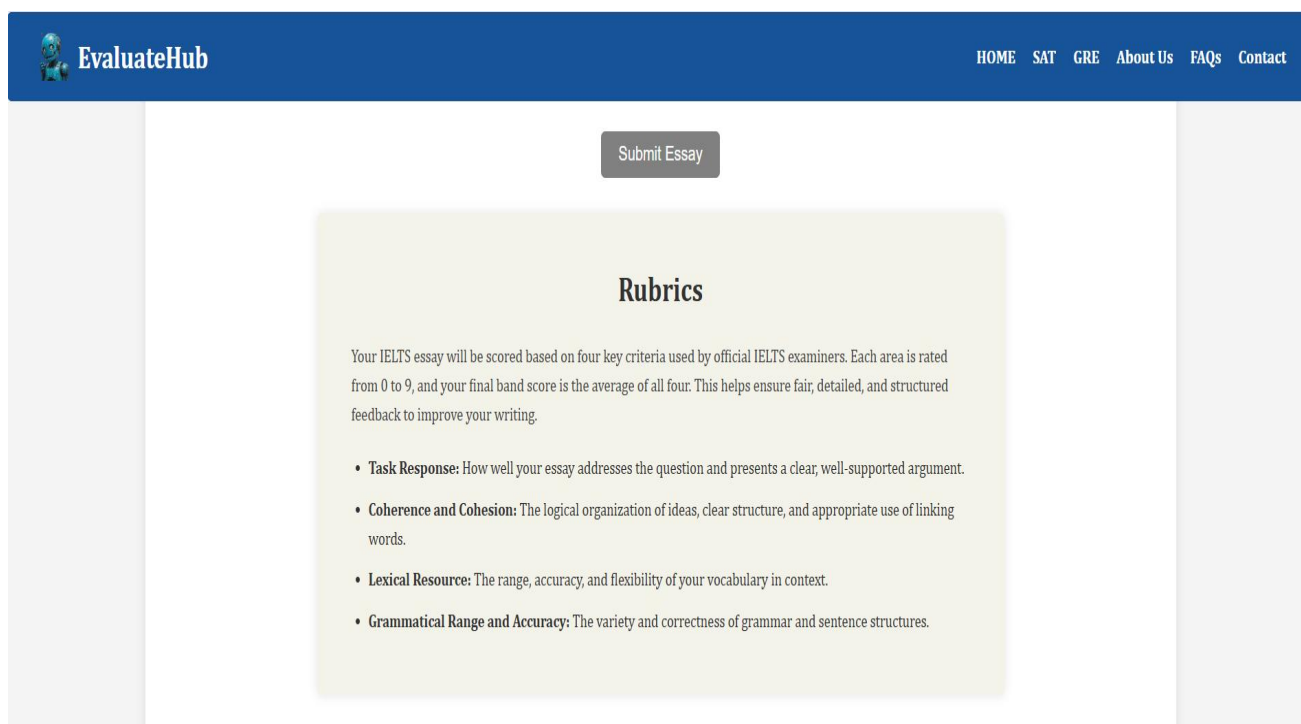


Fig 3.8 Evaluation Result with Feedback

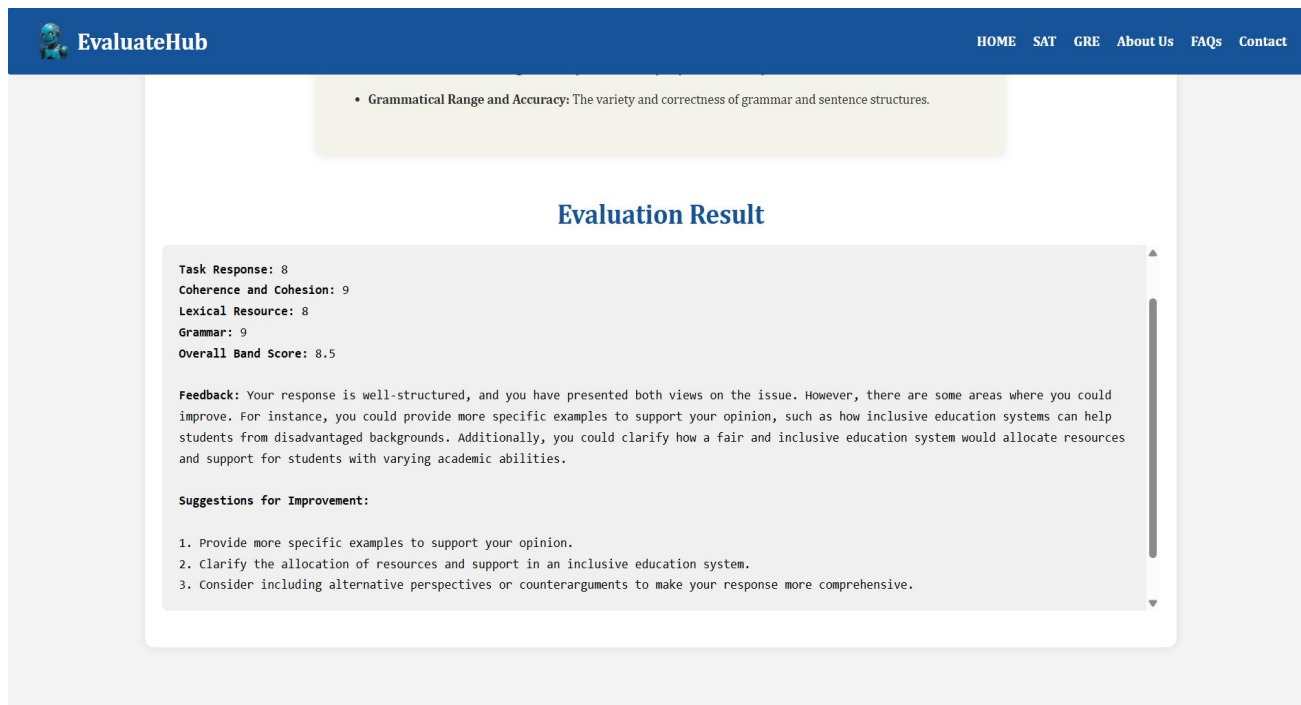



Fig 3.9: IELTS Scoring Rubrics Display

 EvaluateHub

HOME IELTS GRE About Us FAQs Contact

SAT Essay Evaluation

Essay Prompt

Do you agree or disagree with the following statement? "Schools should require students to study a foreign language." Use specific reasons and examples to support your position.

Time: 48:11 Word Count: 400 Minimum: 400 words


I strongly agree with the statement that schools should require students to study a foreign language. In today's increasingly globalized world, being able to communicate in more than one language offers significant academic, professional, and cultural benefits that can shape a student's future in powerful ways.

First and foremost, learning a foreign language enhances cognitive development. Research shows that bilingual individuals tend to have better memory, problem-solving skills, and multitasking abilities. Studying another language also improves students' understanding of grammar and vocabulary in their native language. This can lead to improved performance in other academic areas, such as reading comprehension and writing.

Moreover, being proficient in a foreign language can significantly boost career opportunities. In a global economy, companies often seek employees who can communicate with international clients or partners. For example, someone fluent in Spanish, Mandarin, or French may have more job options in fields like international business, diplomacy, tourism, or even healthcare. Requiring language study in school ensures that all students are given

Submit Essay

Fig 3.10 SAT Essay Submission Interface

 EvaluateHub

HOME IELTS GRE About Us FAQs Contact

other academic areas, such as reading comprehension and writing.

Moreover, being proficient in a foreign language can significantly boost career opportunities. In a global economy, companies often seek employees who can communicate with international clients or partners. For example, someone fluent in Spanish, Mandarin, or French may have more job options in fields like international business, diplomacy, tourism, or even healthcare. Requiring language study in school ensures that all students are given

Submit Essay

Rubrics

Your SAT essay will be evaluated based on five essential criteria. These are designed to reflect the standards of academic writing expected in college-level tests and practice exams. Each essay receives a single writing score between 1 and 8 based on the following elements:

- **Command of Evidence:** The relevance and strength of the examples and references used to support your analysis.
- **Focus and Coherence:** The clarity of your central idea and the logical flow and organization of your essay.
- **Style and Formal Tone:** Use of academic language, formal tone, and appropriate word choices without slang.
- **Grammar and Usage:** Accuracy in sentence structure, punctuation, and grammar throughout the response.
- **Vocabulary & Sentence Variety:** Use of varied sentence structures and vocabulary that enhances clarity and engagement.

Fig 3.11 Evaluation Result with Feedback

Evaluation Result

Thesis: 9

Support: 9

Organization: 8


Evidence: 9

Language Use: 8

Total Score: 86

Feedback: Your essay provides clear and concise reasons for why schools should require students to study a foreign language. You offer several academic, professional, and cultural benefits that can result from language learning. However, there are some areas where you could improve your argumentation. For instance, you could provide more specific examples of how language learning can enhance cognitive development, such as improved memory or problem-solving skills. Additionally, you could address potential criticisms of foreign language education, such as the argument that it is too difficult or not immediately useful.

Fig 3.12 SAT Scoring Rubrics Display

 EvaluateHub

HOME IELTS GRE About Us FAQs Contact

Evaluation Result

Suggestions for Improvement:

1. Provide more concrete examples of how language learning can benefit cognitive development.
2. Address potential criticisms of foreign language education and provide counterarguments to these criticisms.
3. Consider exploring the cultural benefits of language learning in greater detail, such as how it can promote cross-cultural understanding and appreciation.

Overall, your essay presents a strong argument for why schools should require students to study a foreign language. With some additional support and refinement, your argument could be even more compelling.

Future Directions

■ Reading Part Evaluation

🔊 Listening Part Evaluation

🎤 Speaking Part Evaluation

Our Research

Logistic Regression | Random Forest

Support Vector Machines | LSTM

Transformers (BERT, RoBERTa) | DeepSeek

Technologies Used

Django | Node.js | Python

HTML | CSS | JavaScript

LLaMA 2 | Render & Netlify

Developers

Aleena Sajid

Sawaira Iqbal

Sameer Faisal

© 2025 EvaluateHub — Contact: im.aleena.sajid@gmail.com | 0324-6575243

Fig 3.13 SAT essay suggestion and improvement

4. System Design

4.1 System Architecture Diagram

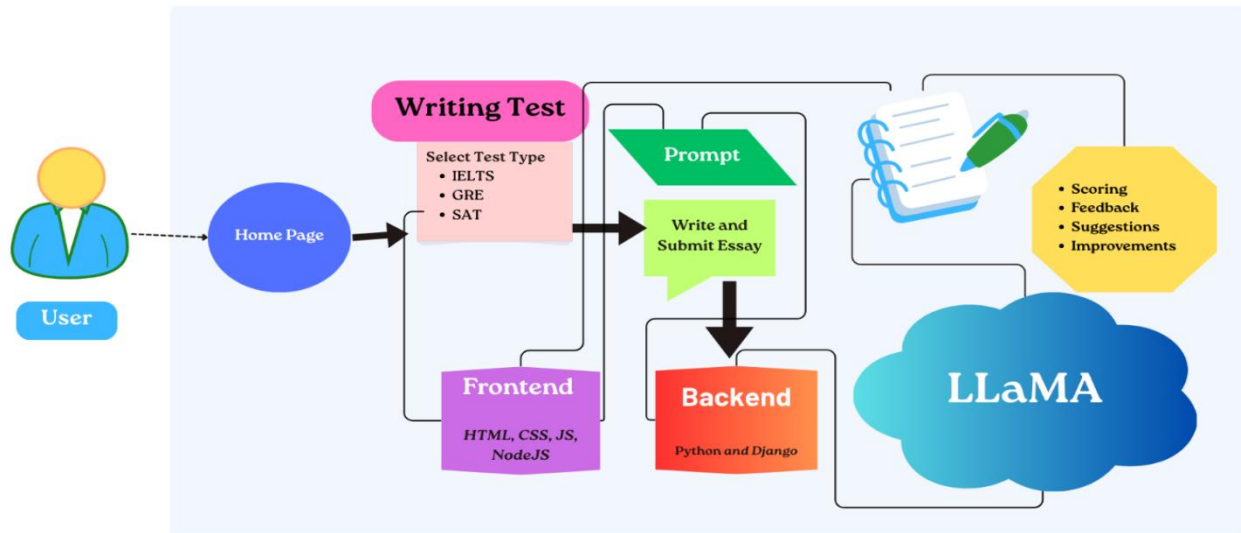


Fig 4.1 System Architecture

4.2 Class Diagram

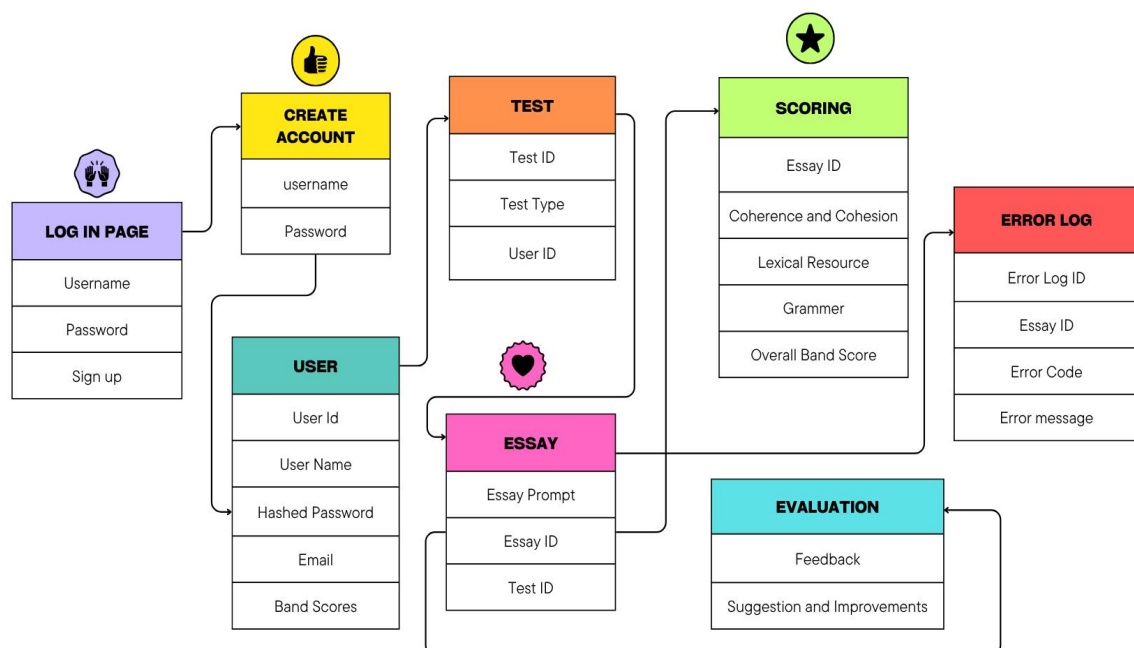


Fig 4.2 System Class diagram

4.3 Sequence Diagram

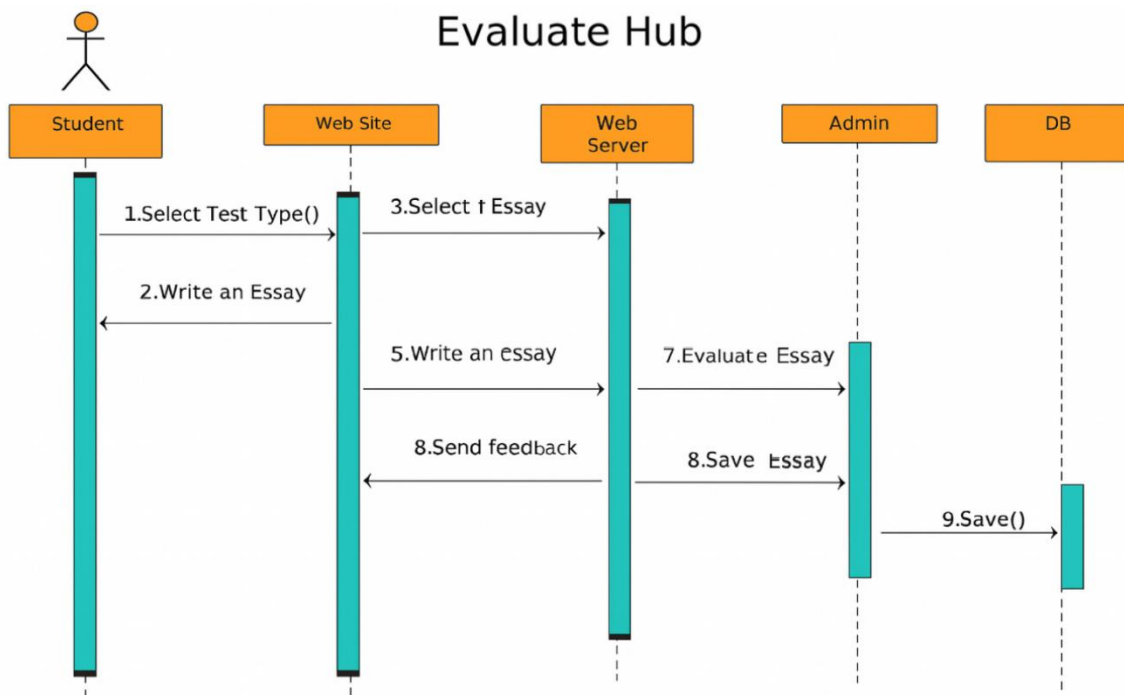


Fig 4.3 System Sequence Diagram

4.4 System Collaboration Diagram

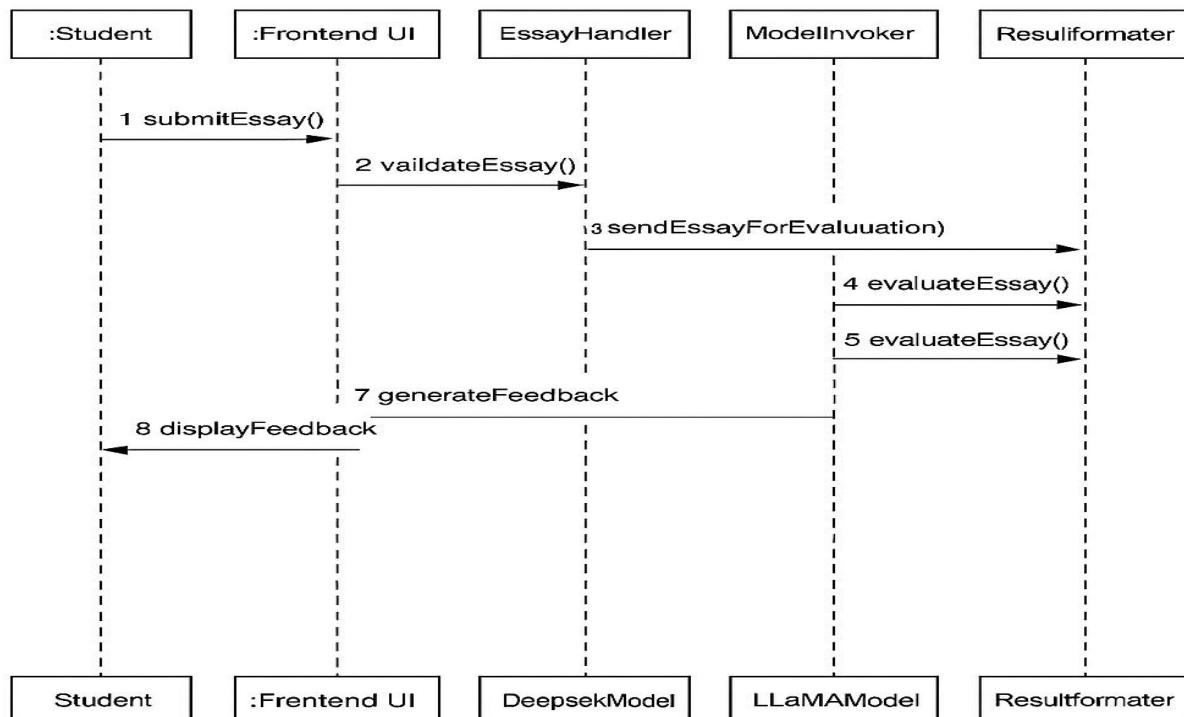


Figure 4.4 System Collaboration Diagram

4.5 System Flow

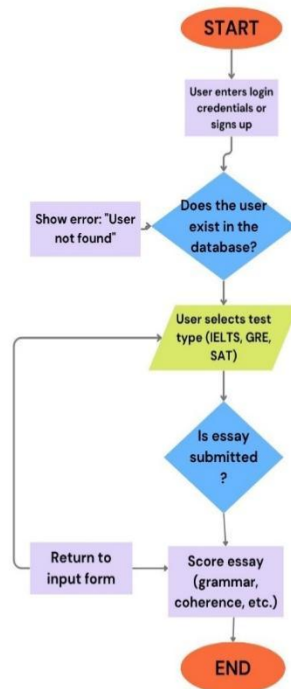


Fig 4.5 System Flow

4.6 Data Dictionary

A data dictionary lists and explains all variables used in the essay evaluation system.

Table 4.1: Data Dictionary

Element Name	Type	Validation	Mandatory	Remarks
user_id	Integer	Auto-increment, Positive Number	Yes	Unique primary key to identify each user
full_name	Varchar(100)	Alphabets and spaces only	Yes	Stores the user's full name
email	Varchar(100)	Valid email format	Yes	Used for login and communication
password	Varchar(255)	Minimum 8 characters, hashed using Bcrypt	Yes	Secured login credential
essay_id	Integer	Auto-increment, Positive Number	Yes	Unique identifier for each essay

test_type	Varchar(50)	IELTS / GRE / SAT	Yes	Specifies which test the essay belongs to
essay_text	Text	Not Empty	Yes	Stores the full written essay
submission_date	DateTime	System-generated	Yes	Tracks when the essay was submitted
grammar_score	Float	Range 0–9	Yes	Score given by the model for grammar
coherence_score	Float	Range 0–9	Yes	Score for logical flow and structure
vocabulary_score	Float	Range 0–9	Yes	Evaluates word usage and diversity
final_band_score	Float	Range 0–9	Yes	Aggregated score generated by LLaMA or DeepSeek
feedback_text	Text	System-generated	Optional	Detailed feedback/suggestions for the user

5. Implementation details

The system uses a LLaMA model integrated with a web interface to evaluate essays under word and time constraints.

5.1 Development Setup

The development of Evaluate Hub involved a combination of modern AI libraries and web technologies. Below is a list of tools and their roles in the system:

Table 4.2: Tool and Technologies

Tool / Technology	Purpose
Python 3.10	Core programming language for backend development and model scripting
TensorFlow / Transformers	Training and using deep learning and LLM models (e.g., LLaMA, BERT, DeepSeek)
Pandas, NumPy	Data loading, preprocessing, and matrix operations
NLTK	Natural language processing tasks like tokenization and stop word removal
Matplotlib, Seaborn	Visualizing model accuracy and feature analysis

Jupyter Notebook / Google Colab	Experimentation and model benchmarking environment
Node.js / React.js	Frontend and backend for web platform
Postman	API testing for essay evaluation and model scoring endpoints
GitHub	Version control and collaborative code development
Render / Localhost	Web app deployment and local testing
VS Code	Primary code editor

5.2 Deployment setup

This was carried out in 2 key environments i.e. local (dev and test) and cloud-based (public demo and multi-factor testing).

Local Deployment:

- **Platform:** Localhost / XAMPP as the platform in the database and backstage of the logic.
- **Frontend:** React.js is being delivered by development server.
- **Model Hosting:** Python models (LLaMA, DeepSeek API) are deployed and host locally using Flask based microservices or Jupyter Notebook endpoints.
- **Database:** MySQL coupled with PHP so as to access the admin panel and store the information provided.
- **Use Case:** It is used when integrating and debugging in early stages of modeling.

5.3 Algorithms

While the entire codebase is extensive, the core functionality relies on the following important custom or improved algorithms:

- **Prompt Generation Algorithm:**
A dynamic message is generated and sent to the AI model (LLaMA or DeepSeek) based on the selected test type (IELTS, SAT, GRE). The prompt includes specific scoring rubrics and a request for structured feedback and suggestions.
- **Essay Evaluation Handling Algorithm:**
A custom handler receives the essay, selects the appropriate evaluation model (LLaMA for IELTS, DeepSeek for SAT/GRE), sends the prompt, receives the JSON response, and extracts the score and feedback cleanly.
- **Rubric-Based Output Mapping Algorithm:**
This algorithm maps the model's raw output to the predefined rubrics and converts it into a structured format (Task Response, Coherence, etc., for IELTS) for display on the frontend.

5.4 Constraints

The system enforces word count and time limits to simulate real standardized test conditions.

5.4.1 Assumptions

- We assume the LLaMA and DeepSeek model will remain available via the Ollama platform without major architectural changes.
- We assume users will have a stable internet connection to interact with the system.
- We assume only writing (essay) input will be submitted no images or attachments.

5.4.2 System Constraints

- The system must operate with both Django and Node.js servers running simultaneously.
- The LLaMA model must be locally hosted and running using the Ollama environment.
- Essays must be in plain text format; PDF or docx uploads are not currently supported.
- Evaluation must occur sequentially batch processing is not implemented.

5.4.3 Restrictions

- The client specified that only the writing (essay) module should be implemented at this stage.
- No integration with third-party storage (e.g., cloud database or cloud file storage) was allowed.
- External AI models like GPT-4 were not permitted due to budget or access limitations.

5.4.4 Limitations

- The system does **not** currently support **reading, listening, or speaking** evaluation.
- Voice or audio input is not supported.
- No plagiarism detection is included.
- The system cannot detect or penalize off-topic essays it only evaluates based on content quality.

6. Testing

The system was tested using sample essays to validate scoring accuracy and performance under constraints.

6.1 Extended Test Cases

Essay Prompt: "Some people think that governments should give financial support to creative artists such as painters and musicians. Others believe that artists should be funded by alternative sources. Discuss both views and give your own opinion."

There are differing views about whether governments should financially support creative artists. While some argue that such funding is a waste of public money, others believe that artists play a vital role in society and deserve governmental

assistance.

On the one hand, those who oppose government funding for artists often claim that there are more urgent priorities, such as education, healthcare, or infrastructure. According to this view, artists should rely on alternative sources such as private sponsorships, art institutions, or the open market to sustain their careers. In many cases, popular musicians and painters already earn significant amounts through exhibitions and performances.

On the other hand, proponents of public funding argue that art contributes to the cultural and intellectual development of a country. Many talented artists may not gain recognition in the competitive private sector and could be overlooked without state support. Government grants can ensure that all forms of art not just commercially successful ones survive and flourish.

In my opinion, a balanced approach is best. While the government should not carry the entire burden, it should support emerging and struggling artists who offer cultural value but lack financial means. At the same time, established artists can rely on private or commercial support.

In conclusion, although there are valid reasons for both sides, a hybrid model of funding would preserve artistic diversity while being financially sustainable.

Table 6.1: Test case 1

Test Case ID:- 1			Test Design By: Aleena Sajid			
Test Module Name: IELTS Essay Evaluation:			Test Design Date: 03/07/2025			
Test Priority:- High			Test Executed By: Aleena Sajid			
Test Title/Name: Submit Valid IELTS Essay			Test Executed Date: 03/07/2025			
Description: Validate end-to-end IELTS essay submission and scoring						
Pre-Condition: User is logged in and selects IELTS before submitting						
Dependencies: Backend must be connected and LLM model running						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Navigate to essay page and select IELTS	Test Type: IELTS	IELTS option selected successfully	IELTS selected	Pass	
2	Enter valid essay and click Submit		Essay is accepted and sent to backend	Pass		
3	Receive and display results	Expected: Scores for all metrics (0–9), feedback, and suggestions				
Post Condition: User sees the full IELTS evaluation with scores (0–9), feedback, and suggestions for improvement.						

Table 6.2: Test case 2

Test Case ID:- 2			Test Design By: Aleena Sajid			
Test Module Name: IELTS Essay Evaluation			Test Design Date: 03/07/2025			
Test Priority:- Medium			Test Executed By: Aleena Sajid			
Test Title/Name: Submit Empty IELTS Essay			Test Executed Date: 03/07/2025			
Description: Test system response when essay field is empty						
Pre-Condition: User selects IELTS and clicks submit without typing essay						
Dependencies: Frontend and backend validation must be enabled						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Go to Essay Submission Page		Page loads	Page loads	Pass	
2	Select IELTS test type	Test Type: IELTS	IELTS selected	IELTS selected	Pass	
3	Leave essay field empty	Essay: ""	Error message: "Essay cannot be empty"	Error displayed	Pass	
Post Condition: Essay is not submitted; user sees an error.						

Table 6.3: Test case 3

Test Case ID:- 3	Test Design By: Aleena Sajid
Test Module Name: IELTS Essay Evaluation	Test Design Date: 03/07/2025
Test Priority:- Medium	Test Executed By: Aleena Sajid
Test Title/Name: Submit Essay with Invalid Characters	Test Executed Date: 03/07/2025
Description: Check if essay containing special/invalid characters is rejected	
Pre-Condition: User selects IELTS and inputs an essay with garbage content	
Dependencies:	

sStep	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Navigate to submission page		Page opens	Page opens	Pass	
2	Select IELTS	Test Type: IELTS	IELTS selected	IELTS selected	Pass	
3	Enter invalid characters	Essay: "@#\$%^&**()&^%\$S\$#"	Error message: "Invalid essay format"	Error shown	Pass	
Post Condition: Essay not submitted.						

Test Case Essay ID: 4

Essay: Technology has become an inseparable part of modern human life. From communication and education to medicine, transportation, and entertainment, technology has transformed how people live, work, and interact. While some individuals argue that these technological advancements have made life easier, others contend that they have introduced new complications and challenges. This essay will discuss both perspectives before concluding with a personal viewpoint.

On one hand, it is undeniable that technology has made many aspects of life significantly easier and more efficient. One of the clearest examples of this is communication. In the past, sending a message to someone across the world might have taken days or even weeks. Today, with a few taps on a smartphone, people can send instant messages, make video calls, or even participate in virtual meetings regardless of geographical location. This convenience has connected people from different cultures and created a more globalized world.

In the field of education, technology has removed many traditional barriers. E-learning platforms, digital textbooks, and online courses have made education more accessible to people in remote or underserved regions. Students can learn at their own pace, revisit lectures, and explore subjects beyond their school curriculum. During the COVID-19 pandemic, millions of students continued their education thanks to online tools. Without technology, educational continuity would not have been possible. Moreover, technology has played a revolutionary role in healthcare. Advanced imaging technologies, robotic surgeries, wearable health monitors, and telemedicine have drastically improved patient care. Early diagnosis and real-time monitoring of health conditions have saved countless lives. In rural areas, doctors can consult patients remotely, allowing better treatment for those who might otherwise go without care.

Similarly, in daily life, technological advancements have contributed to time-saving and convenience. Home automation systems, smart appliances, online shopping, GPS navigation, and ride-hailing apps have streamlined everyday tasks. Tasks that once required physical improving productivity. In professional settings, automation has boosted output, reduced errors, and freed up human workers from repetitive tasks.

On the other hand, critics of technology argue that while it has made some tasks easier, it has also made life more complicated and stressful. One major concern is the loss of privacy. With the widespread use of smartphones, social media, and smart devices, personal data is constantly being collected and stored. People are often unaware of how their data is being used, and data breaches or misuse of information can have serious consequences.

Moreover, technology has led to an "always online" culture, where people are expected to be reachable and responsive at all times. Work-life balance has eroded, with emails and work notifications extending into evenings and weekends. Employees feel pressured to remain constantly connected, leading to stress, burnout, and decreased quality of personal relationships.

Another complication arises from the over-dependence on technology. Many people, especially younger generations, rely heavily on digital devices for even simple tasks. For example, the widespread use of navigation apps has reduced people's natural sense of direction. Overreliance on spell checkers, calculators, or AI tools may hinder cognitive development and

problem-solving skills. In extreme cases, addiction to video games or social media can negatively impact mental health and social interaction.

Technology has also contributed to unemployment in certain sectors. With the rise of automation and artificial intelligence, many low-skilled jobs have been replaced by machines. This has created economic disparity and social unrest in some areas. While new tech-related jobs are created, they often require specialized skills that displaced workers may not possess, further widening the skill gap.

Additionally, technological development can deepen inequality between developed and developing nations. While high-income countries benefit from cutting-edge tools and infrastructure, poorer regions may lack access to even basic technologies. This digital divide prevents equal participation in education, healthcare, and economic opportunities.

Furthermore, environmental concerns have been raised regarding the production and disposal of electronic devices. E-waste is a growing problem, and the mining of rare earth elements used in electronics harms ecosystems. Despite the push for green technology, the energy demands of data centers, cryptocurrency mining, and large-scale tech infrastructure contribute to carbon emissions.

In terms of human interaction, technology has made communication easier but arguably less meaningful. People may spend hours on their phones without speaking to the person next to them. Social media fosters comparison, envy, and low self-esteem. Online platforms can also be breeding grounds for misinformation, cyberbullying, and echo chambers that divide societies rather than unify them.

Despite these valid concerns, it is important to recognize that technology itself is neutral — its impact depends on how it is used. The solution lies in developing a responsible and balanced approach to its integration in life. Governments should establish policies to protect digital privacy, regulate tech companies, and ensure access to technology for all socioeconomic groups. Individuals must also take responsibility for managing screen time, using tools mindfully, and staying informed.

In my opinion, technology has overall made life easier, but only for those who are able to adapt, use it wisely, and have access to it. While the problems it introduces cannot be ignored, most of them stem from misuse, overdependence, or lack of regulation rather than the technology itself. The benefits in healthcare, education, connectivity, and convenience outweigh the complications, provided society learns to manage the side effects.

In conclusion, while technology has undeniably made certain aspects of life more complicated especially regarding privacy, mental health, and employment it has also provided powerful tools that improve efficiency, safety, and global connection. Striking a balance between usage and awareness is crucial to ensuring that technology serves humanity positively rather than dominates it.

Table 6.4: Test case 4

Test Case ID:- 4	Test Design By: Aleena Sajid
Test Module Name: IELTS Essay Evaluation	Test Design Date: 03/07/2025
Test Priority:- Medium	Test Executed By: Aleena Sajid
Test Title/Name: Submit Essay Longer Than Max Limit	Test Executed Date: 03/07/2025
Description: Validate essay length check	
Pre-Condition: User selects IELTS and submits an essay >1000 words	
Dependencies: Backend must enforce word limits	

Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Navigate to page		Page loads	Page loads	Pass	
2	Select IELTS	IELTS	Selected	Selected	Pass	
3	Input long essay		Error: "Essay too long"	Error shown	Pass	
Post Condition: User sees validation error.						

Test Case Essay ID: 5

Essay: University education is essential for personal and national growth. Some people believe it should be free for all. I partly agree with this view. Free education would help poor students pursue their dreams and reduce social inequality. However, making it completely free for everyone may not be financially possible for many governments. It could lead to over-enrollment and strain resources. A better solution is to offer free education to those in need, while others pay a reasonable fee. This way, access is fair, and the system remains sustainable.

Table 6.5: Test case 5

Test Case ID:- 5			Test Design By: Aleena Sajid			
Test Module Name: IELTS Essay Evaluation			Test Design Date: 03/07/2025			
Test Priority:- High			Test Executed By: Aleena Sajid			
Test Title/Name: Backend Error on Submit			Test Executed Date: 03/07/2025			
Description: Check system behavior when backend times out or fails						
Pre-Condition: Backend is intentionally turned off or returns 500 error						
Dependencies: Backend must simulate failure or timeout						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Open essay submission page and select IELTS	Test Type: IELTS	Page and dropdown load correctly	Pass	Pass	
2	Enter valid essay and click Submit		Request is sent to backend	Pass	Pass	
3	Simulate backend timeout	Server down or slow	Error: "Server not responding"	Pass	Pass	
Post Condition: Essay is not submitted , and the user sees a clear error message indicating a backend/server issue. No scores or feedback are received or stored.						

Test Case Essay ID: 6

Essay: University education is essential for personal and national growth. Some people believe it should be free for all. I partly agree with this view. Free education would help poor students pursue their dreams and reduce social inequality. However, making it completely free for everyone may not be financially possible for many governments. It could lead to over-enrollment and strain resources. A better solution is to offer free education to those in need, while others pay a reasonable fee. This way, access is fair, and the system remains sustainable.

Table 6.6: Test case 6

Test Case ID:- 6			Test Design By: Aleena Sajid			
Test Module Name: IELTS Essay Evaluation			Test Design Date: 03/07/2025			
Test Priority:- High			Test Executed By: Aleena Sajid			
Test Title/Name: Handle Invalid Score Returned			Test Executed Date: 03/07/2025			
Description: Ensure scores outside valid IELTS range (0–9) are handled						
Pre-Condition: Backend returns invalid score (e.g., 12)						
Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Submit a valid IELTS essay		Essay accepted for processing	Pass	Pass	
2	Backend returns invalid score	Task Response = 12	System identifies score as invalid	Pass	Pass	
3	Display result		Score not shown or defaulted to “N/A”	Pass	Pass	
Post Condition: The system rejects or ignores invalid score data (e.g., score > 9). No results are displayed to the user. The issue may be logged for developer review.						

Test Case Essay ID: 7

Essay: University education is essential for personal and national growth. Some people believe it should be free for all. I partly agree with this view. Free education would help poor students pursue their dreams and reduce social inequality. However, making it completely free for everyone may not be financially possible for many governments. It could lead to

over-enrollment and strain resources. A better solution is to offer free education to those in need, while others pay a reasonable fee. This way, access is fair, and the system remains sustainable.

Table 6.7: Test case 7

Test Case ID:- 7			Test Design By: Aleena Sajid			
Test Module Name: IELTS Essay Evaluation			Test Design Date: 03/07/2025			
Test Priority:- High			Test Executed By: Aleena Sajid			
Test Title/Name: Ensure Feedback and Suggestions are Shown			Test Executed Date: 03/07/2025			
Description: Ensure feedback and improvement suggestions are displayed						
Pre-Condition: Valid essay submitted and evaluated						
Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Select IELTS and enter valid essay		Essay accepted	Pass	Pass	
2	Submit essay and receive evaluation		Feedback and suggestions included in response	Pass	Pass	
3	Display on frontend		Scores, feedback and suggestions shown	Pass	Pass	
Post Condition: The user sees the complete result including scores (0–9), written feedback , and suggestions clearly displayed on the result screen. Data may be stored for later access.						

Test Case Essay ID: 8

Essay: In the passage, the author argues that public libraries remain a vital part of society despite the rise of digital technology. To support this claim, the author uses emotional appeals, facts, and expert opinions.

The author begins by appealing to the reader’s sense of nostalgia, describing the library as “a safe place where learning meets community.” This emotional connection makes the reader more likely to support the preservation of libraries. Additionally, the author strengthens the argument by providing statistical evidence citing a study showing that over 60% of people still use libraries regularly for free access to books, internet, and job resources.

To further build credibility, the author quotes respected educators and urban planners who emphasize the role libraries play in bridging the digital divide. These experts argue that not everyone can afford digital devices or fast internet, making public libraries essential.

Through a mix of emotional language, factual support, and credible expert voices, the author builds a convincing case for maintaining public libraries. The logical structure and balanced reasoning help persuade readers of the ongoing value of these institutions.

Table 6.8: Test case 8

Test Case ID:- 8			Test Design By: Aleena Sajid			
Test Module Name: SAT Essay Evaluation			Test Design Date: 03/07/2025			
Test Priority:- High			Test Executed By: Aleena Sajid			
Test Title/Name: Submit Valid SAT Essay			Test Executed Date: 03/07/2025			
Description: Validate full SAT essay submission and scoring process						
Pre-Condition: User selects SAT and provides an appropriate essay						
Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Navigate to essay page and select SAT	Test Type: SAT	SAT selected successfully	SAT selected successfully	Pass	
2	Enter and submit valid essay		Essay submitted to backend	Essay submitted to backend	Pass	
3	Receive and display result		Scores and feedback shown on UI	Scores and feedback shown on UI	Pass	
Post Condition: SAT scores and feedback are displayed to the user and optionally stored for future access.						

Table 6.9: Test case 9

Test Case ID:- 9			Test Design By: Aleena Sajid			
Test Module Name: SAT Essay Evaluation			Test Design Date: 03/07/2025			
Test Priority:- High			Test Executed By: Aleena Sajid			

Test Title/Name: Reject Empty or Too Short Essay			Test Executed Date: 03/07/2025			
Description: Ensure that essays with no content or insufficient word count are rejected						
Pre-Condition: User selects SAT and submits a blank or very short essay						
Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Select SAT and navigate to essay submission page	Test Type: SAT	SAT selected successfully	SAT selected successfully	Pass	
2	Submit blank or short essay	Essay:'''	System blocks submission	System blocks submission	Pass	
3	Observe UI message		Error: "Essay too short. Please write at least 200 words."	Error: "Essay too short. Please write at least 200 words."	Pass	
Post Condition: Essay is not submitted , and user is clearly informed about the minimum length requirement (200 words).						

Test Case Essay ID: 10

Essay: In the passage, the author argues that public libraries remain a vital part of society despite the rise of digital technology. To support this claim, the author uses emotional appeals, facts, and expert opinions.

The author begins by appealing to the reader's sense of nostalgia, describing the library as "a safe place where learning meets community." This emotional connection makes the reader more likely to support the preservation of libraries. Additionally, the author strengthens the argument by providing statistical evidence citing a study showing that over 60% of people still use libraries regularly for free access to books, internet, and job resources.

To further build credibility, the author quotes respected educators and urban planners who emphasize the role libraries play in bridging the digital divide. These experts argue that not everyone can afford digital devices or fast internet, making public libraries essential.

Through a mix of emotional language, factual support, and credible expert voices, the author builds a convincing case for maintaining public libraries. The logical structure and balanced reasoning help persuade readers of the ongoing value of these institutions.

Table 6.10: Test case 10

Test Case ID:- 10			Test Design By: Aleena Sajid			
Test Module Name: SAT Essay Evaluation			Test Design Date: 03/07/2025			
Test Priority:- Medium			Test Executed By: Aleena Sajid			
Test Title/Name: Wrong Test Type Selected			Test Executed Date: 03/07/2025			
Description: Ensure SAT essay is not evaluated under IELTS or GRE by mistake						
Pre-Condition: User writes a SAT-style essay but selects the wrong test type						
Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Navigate to essay page and select IELTS instead of SAT	Test Type: IELTS	Test type selected as IELTS	Test type selected as IELTS	Pass	
2	Paste SAT-style essay and click submit		System detects mismatch	System detects mismatch	Pass	
3	Observe feedback or error		Essay format does not match selected test type.” or poor score returned	Essay format does not match selected test type.” or poor score returned	Pass	
Post Condition: Essay is evaluated inaccurately, or the system returns a validation warning or incorrect score alert to avoid misclassification.						

Test Case Essay ID: 11

Essay: In the article, the author presents a compelling argument advocating for the preservation and expansion of art programs in schools. Arguing that art education is essential to student development, the author appeals to both logic and emotion while supporting claims with credible evidence, structured reasoning, and rhetorical strategies. Through these techniques, the author effectively persuades the reader to recognize the value of arts in academic settings.

To begin with, the author introduces the argument by highlighting a widely acknowledged problem: declining support for the arts in education. Budget cuts, policy shifts, and a greater focus on STEM subjects have led many schools to reduce or eliminate visual art, music, and drama programs. By starting with this contextual framing, the author sets the stage for urgency and relevance an important rhetorical move that ensures the reader understands the stakes of the issue.

The author then uses **statistical evidence** to strengthen the credibility of the argument. For instance, the article cites a nationwide study from the National Endowment for the Arts, which reports that students involved in school art programs score higher on standardized tests and are more likely to attend college. This quantitative data supports the author’s central claim with objective numbers, making the argument harder to dismiss as opinion-based. Additionally, such evidence reassures skeptics who prioritize academic outcomes over creativity.

Another notable strategy is the use of **expert opinion**. The author includes insights from Dr. Angela Brookstone, a respected child psychologist, who asserts that “creative expression is directly linked to emotional intelligence and

problem-solving ability.” This expert backing serves a dual purpose: it appeals to ethos (credibility) while also reinforcing the emotional and cognitive benefits of art. Including such a source allows the author to shift the reader's view of art from a ‘nice-to-have’ to a developmental necessity.

In addition to citing evidence, the author also applies **logical reasoning** to explain why the removal of art programs is short-sighted. The article explains that although schools may save money by cutting art classes, this often results in long-term educational harm. Students who lack creative outlets tend to disengage from school, which can lead to increased dropout rates and behavioral issues. The author logically connects the presence of art to increased engagement, reduced stress, and stronger classroom communities, thus painting a picture of art as an investment rather than an expense.

To appeal to the reader’s **sense of morality and fairness**, the author uses a rhetorical question: “Shouldn’t every child, regardless of background, have the opportunity to explore their creative potential?” This question compels readers to reflect on equity in education. By framing the issue as a matter of opportunity and justice, the author stirs the reader’s conscience. The argument becomes not just about policy but about what society values and for whom.

The author further enhances the argument through the use of **anecdotal evidence**. A particularly moving paragraph describes a student named Luis, who struggled with academic subjects but discovered a passion and talent for painting in high school. Through his involvement in art classes, Luis developed confidence, improved his overall grades, and eventually received a scholarship to an art college. This real-life story adds a human face to the broader argument. It creates emotional resonance, making the data and logic more relatable and memorable.

Moreover, the author preemptively **addresses counterarguments**, a powerful persuasive tactic. Recognizing that some readers may argue that art is not as critical as math or science, the author counters by noting that many top-performing countries in education such as Finland and South Korea integrate art deeply into their curriculums. This shows that a well-rounded education does not come at the expense of academic rigor; rather, the arts complement and enhance cognitive abilities. By acknowledging and rebutting opposing views, the author strengthens the overall persuasiveness of the piece.

Additionally, **language choice and tone** play subtle but important roles. The author maintains a tone that is passionate but respectful, assertive but not aggressive. Words like “vital,” “nurture,” “expression,” and “potential” are used throughout the essay to evoke a sense of value and growth. These word choices align with the overall argument — that art nurtures not just skills but also the human spirit.

The conclusion ties the entire argument together by calling for action urging school administrators, policymakers, and parents to advocate for robust art programs. The author reinforces earlier points and ends with a quote from Leonardo da Vinci: “Art is the queen of all sciences communicating knowledge to all the generations of the world.” This final quote resonates with the reader and leaves a lasting impression, reminding us of the historical and cultural importance of creativity.

In summary, the author builds a convincing and passionate argument in favor of art education through the effective use of evidence, logical reasoning, rhetorical devices, and emotional appeal. By blending expert opinion, statistics, personal anecdotes, and counterargument refutation, the author successfully addresses various types of readers — from logical thinkers to emotionally driven advocates. The reader is left with a clear understanding of why arts should not just survive but thrive within the academic system.

Table 6.11: Test case 11

Test Case ID:- 11	Test Design By: Aleena Sajid
Test Module Name: SAT Essay Evaluation	Test Design Date: 03/07/2025
Test Priority:- Medium	Test Executed By: Aleena Sajid
Test Title/Name: ubmit Over-Length SAT Essay	Test Executed Date: 03/07/2025
Description: Check how the system handles essays longer than the accepted word count	
Pre-Condition: User selects SAT and submits an essay exceeding 1000 words	
Dependencies:	

Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Select SAT and open submission page	Test Type: SAT	SAT selected successfully	SAT selected successfully	Pass	
2	Submit long essay >1000 words		System detects length overflow	System detects length overflow	Pass	
3	View validation result		Error message: "Essay too long. Please limit to 650 words."	Error message: "Essay too long. Please limit to 650 words."	Pass	
Post Condition: The essay is not evaluated, and a length validation error is shown to the user, enforcing SAT's expected range (typically 400–650 words).						

Test Case Essay ID: 12

passage, the author argues that public libraries remain a vital part of society despite the rise of digital technology. To support this claim, the author uses emotional appeals, facts, and expert opinions.

The author begins by appealing to the reader's sense of nostalgia, describing the library as "a safe place where learning meets community." This emotional connection makes the reader more likely to support the preservation of libraries. Additionally, the author strengthens the argument by providing statistical evidence citing a study showing that over 60% of people still use libraries regularly for free access to books, internet, and job resources.

To further build credibility, the author quotes respected educators and urban planners who emphasize the role libraries play in bridging the digital divide. These experts argue that not everyone can afford digital devices or fast internet, making public libraries essential.

Through a mix of emotional language, factual support, and credible expert voices, the author builds a convincing case for maintaining public libraries. The logical structure and balanced reasoning help persuade readers of the ongoing value of these institutions.

Table 6.12: Test case 12

Test Case ID:- 12	Test Design By: Aleena Sajid
Test Module Name: SAT Essay Evaluation	Test Design Date: 03/07/2025
Test Priority:- High	Test Executed By: Aleena Sajid
Test Title/Name: Backend Timeout/Error on Submit	Test Executed Date: 03/07/2025
Description: Ensure system handles backend failure or timeout gracefully during SAT essay evaluation	
Pre-Condition: User selects SAT and writes a valid essay; backend service is unavailable	

Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Select SAT and write valid essay		Essay input accepted	Essay input accepted	Pass	
2	Submit while backend is down	Backend unavailable or unresponsive	Error triggered: "Server not responding"	Error triggered: "Server not responding"	Pass	
3	Observe UI response		User remains on page with error message displayed	User remains on page with error message displayed	Pass	
Post Condition: Essay is not evaluated, and the user sees a clear error message about the backend issue, ensuring a graceful fail with no data loss.						

Test Case Essay ID: 13

Essay: In the passage, the author argues that public libraries remain a vital part of society despite the rise of digital technology. To support this claim, the author uses emotional appeals, facts, and expert opinions.

The author begins by appealing to the reader's sense of nostalgia, describing the library as "a safe place where learning meets community." This emotional connection makes the reader more likely to support the preservation of libraries. Additionally, the author strengthens the argument by providing statistical evidence citing a study showing that over 60% of people still use libraries regularly for free access to books, internet, and job resources.

To further build credibility, the author quotes respected educators and urban planners who emphasize the role libraries play in bridging the digital divide. These experts argue that not everyone can afford digital devices or fast internet, making public libraries essential.

Through a mix of emotional language, factual support, and credible expert voices, the author builds a convincing case for maintaining public libraries. The logical structure and balanced reasoning help persuade readers of the ongoing value of these institutions.

Table 6.13: Test case 13

Test Case ID:- 13	Test Design By: Aleena Sajid
Test Module Name: SAT Essay Evaluation	Test Design Date: 03/07/2025
Test Priority:- High	Test Executed By: Aleena Sajid
Test Title/Name: Handle Invalid Score Returned	Test Executed Date: 03/07/2025
Description: Ensure system handles invalid scores (outside 2–8 range) returned from backend	

Pre-Condition: User submits a valid SAT essay; backend returns invalid scores						
Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Select SAT and submit valid essay		Essay accepted	Essay accepted	Pass	
2	Simulate backend response with invalid scores		System flags scores as invalid	System flags scores as invalid	Pass	
3	Observe system behavior		UI shows “Invalid score returned” or skips display	UI shows “Invalid score returned” or skips display	Pass	
Post Condition: The system does not display invalid scores to the user and may show a warning or fallback behavior like “N/A” with error logging.						

Test Case Essay ID: 14

Essay: In the passage, the author argues that public libraries remain a vital part of society despite the rise of digital technology. To support this claim, the author uses emotional appeals, facts, and expert opinions.

The author begins by appealing to the reader’s sense of nostalgia, describing the library as “a safe place where learning meets community.” This emotional connection makes the reader more likely to support the preservation of libraries. Additionally, the author strengthens the argument by providing statistical evidence citing a study showing that over 60% of people still use libraries regularly for free access to books, internet, and job resources.

To further build credibility, the author quotes respected educators and urban planners who emphasize the role libraries play in bridging the digital divide. These experts argue that not everyone can afford digital devices or fast internet, making public libraries essential.

Through a mix of emotional language, factual support, and credible expert voices, the author builds a convincing case for maintaining public libraries. The logical structure and balanced reasoning help persuade readers of the ongoing value of these institutions.

Table 6.14: Test case 14

Test Case ID:- 14	Test Design By: Aleena Sajid
Test Module Name: SAT Essay Evaluation	Test Design Date: 03/07/2025
Test Priority:- High	Test Executed By: Aleena Sajid
Test Title/Name: Display Feedback and Suggestions	Test Executed Date: 03/07/2025

Description: Ensure feedback and suggestions are displayed clearly alongside domain scores						
Pre-Condition: User submits a valid SAT essay; backend returns scores and textual feedback						
Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Select SAT and enter valid essay		Essay accepted	Essay accepted	Pass	
2	Submit essay and receive evaluation		Textual feedback returned from backend	Textual feedback returned from backend	Pass	
3	View results on frontend		Scores + feedback + improvement suggestions displayed	Scores + feedback + improvement suggestions displayed	Pass	
Post Condition: The system displays all scores and feedback to the user in a readable and organized manner						

Test Case Essay ID: 15

Essay: Governments must strike a careful balance between addressing current challenges and preparing for future ones. While immediate problems such as unemployment, poverty, and health crises demand urgent attention, I strongly believe that governments should **not ignore future concerns**. Anticipating and preparing for future problems is essential to ensure long-term stability, security, and sustainability.

First, many of today's most pressing issues are the **result of ignoring future risks in the past**. For instance, climate change now a global crisis could have been mitigated if governments had taken more proactive steps decades ago. Failure to reduce emissions, regulate polluting industries, and promote green energy early on has led to rising sea levels, extreme weather, and environmental disasters that are now harder and more expensive to manage. This example illustrates that **delayed action on future problems often turns them into immediate crises**.

Second, solving current problems without considering the long-term effects can lead to **short-sighted policies**. For example, a government may subsidize fossil fuels to lower energy costs today, but this approach increases pollution and delays investment in renewable technologies. By contrast, policies that are future-focused—such as funding solar and wind energy may not yield instant results but ensure cleaner air, job creation in green sectors, and energy independence in the long run.

Of course, it is true that **some problems require urgent solutions**. A government that fails to address immediate issues like a pandemic, economic recession, or natural disaster can lose public trust and face political instability. However, focusing exclusively on immediate concerns often leads to a **reactive government**, one that is constantly firefighting rather than building a sustainable path forward.

Moreover, future-oriented investments often have **dual benefits** they prepare us for tomorrow and improve lives today. Consider education: investing in better schools, teacher training, and curriculum reform is a future-facing strategy. However, it also improves current learning outcomes, reduces dropout rates, and boosts employment. The same can be said for investments in infrastructure, technology, and healthcare systems.

A well-functioning government must therefore **allocate resources strategically** dedicating sufficient effort to present needs without compromising long-term goals. This doesn't mean future challenges should overshadow present ones, but rather that a **visionary government prepares for both**. By creating long-term plans, funding research and innovation, and involving experts in policymaking, governments can prevent future problems from becoming catastrophes.

In conclusion, while solving immediate problems is essential, **governments must also anticipate and address future challenges**. Only by balancing the present with the future can a nation achieve sustainable progress. A short-term mindset may win political favor, but a long-term vision builds a safer, stronger, and more resilient society.

Table 6.15: Test case 15

Test Case ID:- 15			Test Design By: Aleena Sajid			
Test Module Name: GRE Essay Evaluation			Test Design Date: 03/07/2025			
Test Priority:- High			Test Executed By: Aleena Sajid			
Test Title/Name: Submit Valid GRE Issue Essay			Test Executed Date: 03/07/2025			
Description: Ensure successful submission, scoring, and feedback of a valid Issue essay						
Pre-Condition: User selects GRE > Issue Essay and submits a proper essay						
Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Select GRE > Issue Essay	Test Type: GRE – Issue	Option selected correctly	Option selected correctly	Pass	
2	Paste valid essay (~500 words)		Essay submitted	Essay submitted	Pass	
3	Observe evaluation output		Result displayed properly	Result displayed properly	Pass	
Post Condition: Essay is scored, feedback shown, and all metrics from 0–6 are displayed clearly to the user.						

Test Case Essay ID: 16

Essay: In the passage, the author argues that public libraries remain a vital part of society despite the rise of digital technology. To support this claim, the author uses emotional appeals, facts, and expert opinions.

The author begins by appealing to the reader’s sense of nostalgia, describing the library as “a safe place where learning meets community.” This emotional connection makes the reader more likely to support the preservation of libraries. Additionally, the author strengthens the argument by providing statistical evidence citing a study showing that over 60% of people still use libraries regularly for free access to books, internet, and job resources.

To further build credibility, the author quotes respected educators and urban planners who emphasize the role libraries play in bridging the digital divide. These experts argue that not everyone can afford digital devices or fast internet, making public libraries essential.

Through a mix of emotional language, factual support, and credible expert voices, the author builds a convincing case for maintaining public libraries. The logical structure and balanced reasoning help persuade readers of the ongoing value of these institutions.

Table 6.16: Test case 16

Test Case ID:- 16			Test Design By: Aleena Sajid			
Test Module Name: GRE Essay Evaluation			Test Design Date: 03/07/2025			
Test Priority:- Medium			Test Executed By: Aleena Sajid			
Test Title/Name: Incorrect Essay Type Submission			Test Executed Date: 03/07/2025			
Description: Validate how the system handles when an Argument-style essay is submitted to the Issue section						
Pre-Condition: GRE > Issue Essay selected; user enters Argument-style response						
Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Select GRE > Issue Essay	Test Type: GRE – Issue	Selected correctly	Selected correctly	Pass	
2	Submit Argument-style essay (logic critique)		Warning or poor score	Warning or poor score	Pass	
3	Observe response		Result shows low scores or prompt mismatch alert	Result shows low scores or prompt mismatch alert	Pass	
Post Condition: System either flags style mismatch or assigns low scores due to irrelevant structure.						

Test Case Essay ID: 17

Essay: In the passage, the author argues that public libraries remain a vital part of society despite the rise of digital technology. To support this claim, the author uses emotional appeals, facts, and expert opinions. The author begins by appealing to the reader’s sense of nostalgia, describing the library as “a safe place where learning meets community.” This emotional connection makes the reader more likely to support the preservation of libraries. Additionally, the author strengthens the argument by providing statistical evidence — citing a study showing that over 60% of people still use libraries regularly for free access to books, internet, and job resources.

To further build credibility, the author quotes respected educators and urban planners who emphasize the role libraries play in bridging the digital divide. These experts argue that not everyone can afford digital devices or fast internet, making public libraries essential.

Through a mix of emotional language, factual support, and credible expert voices, the author builds a convincing case for maintaining public libraries. The logical structure and balanced reasoning help persuade readers of the ongoing value of these institutions.

In the passage, the author argues that public libraries remain a vital part of society despite the rise of digital technology. To support this claim, the author uses emotional appeals, facts, and expert opinions.

The author begins by appealing to the reader’s sense of nostalgia, describing the library as “a safe place where learning meets community.” This emotional connection makes the reader more likely to support the preservation of libraries.

Additionally, the author strengthens the argument by providing statistical evidence citing a study showing that over 60% of people still use libraries regularly for free access to books, internet, and job resources. To further build credibility, the author quotes respected educators and urban planners who emphasize the role libraries play in bridging the digital divide. These experts argue that not everyone can afford digital devices or fast internet, making public libraries essential.

Table 6.17: Test case 17

Test Case ID:- 17			Test Design By: Aleena Sajid			
Test Module Name: GRE Essay Evaluation			Test Design Date: 03/07/2025			
Test Priority:- Medium			Test Executed By: Aleena Sajid			
Test Title/Name: Submit Over-Length Essay			Test Executed Date: 03/07/2025			
Description: Check how the system handles essays exceeding GRE's typical length (600 words)						
Pre-Condition: GRE > Argument Essay selected; essay exceeds 1000+ words						
Dependencies:						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Select GRE > Argument Essay	Type: GRE – Argument	Option selected successfully	Option selected successfully	Pass	
2	Submit 1000-word essay		System flags word count exceeded	System flags word count exceeded	Pass	
3	View response		Error: “Essay too long. Please limit to 600 words.”	Error: “Essay too long. Please limit to 600 words.”	Pass	
Post Condition: Essay is not evaluated. Clear message alerts user of maximum word count restriction.						

Test Case Essay ID: 18

Essay: The argument concludes that sales in the southeast region will increase if gourmet baked goods are introduced, just as they did in the northeast region. While this idea may seem reasonable at first glance, the argument is logically flawed and rests on several unsupported assumptions that weaken its persuasiveness.

First, the argument **assumes that the increase in sales in the northeast was solely due to the introduction of gourmet baked goods**. However, correlation does not imply causation. There may have been other factors influencing the sales increase such as better marketing campaigns, seasonal demand, new store openings, changes in pricing, or improvements in customer service. Without clear evidence that baked goods were the primary cause of the increase, the conclusion is questionable.

Second, the argument **assumes that customer preferences in the southeast are similar to those in the northeast**. This overlooks the possibility of regional differences in taste, dietary habits, and cultural food preferences. For instance, customers in the southeast may prefer savory snacks or traditional items over gourmet baked goods. If this is the case, the strategy may not have the same effect and could lead to wasted investment in unpopular products.

Another flawed assumption is that the **economic and competitive environments in the two regions are the same**. The northeast stores may be located in wealthier urban areas with customers willing to pay for high-end pastries, while the southeast stores may be in more price-sensitive markets or areas with greater competition from local bakeries or coffee shops. Without analyzing the demographics and competition in both regions, the argument lacks a solid foundation.

Moreover, the memo does not mention whether the increase in sales in the northeast led to increased **profits**. Gourmet baked goods are often more expensive to source or produce, and if the profit margins are slim or negative, replicating the strategy may not be financially beneficial even if sales volume increases. Profitability, not just sales volume, is key to sustainable business success.

The memo also **fails to consider implementation factors**, such as employee training, inventory management, and supply chain adjustments. Introducing a new product line involves costs and logistical planning. If the southeast stores are not equipped to handle this change effectively, it could lead to operational disruptions and customer dissatisfaction.

In conclusion, the argument that gourmet baked goods will increase sales in the southeast, just as in the northeast, is based on **several unproven assumptions**. To strengthen the argument, the author should provide clear evidence that baked goods caused the increase in the northeast, analyze customer preferences in the southeast, compare economic contexts, and evaluate potential costs. Without addressing these gaps, the recommendation remains unconvincing and risky.

Table 6.18: Test case 18

Test Case ID:- 18			Test Design By: Aleena Sajid							
Test Module Name: GRE Essay Evaluation			Test Design Date: 03/07/2025							
Test Priority:- High			Test Executed By: Aleena Sajid							
Test Title/Name: Handle Backend Timeout/Error			Test Executed Date: 03/07/2025							
Description: Ensure frontend handles server crash or delay during GRE essay submission										
Pre-Condition: User submits essay while evaluation service is down										
Dependencies:										
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes				
1	Select GRE > Argument Essay and enter essay		Essay accepted by UI	Essay accepted by UI	Pass					
2	Submit while backend is down		App shows error gracefully	App shows error gracefully	Pass					
3	Observe UI feedback		User stays on same page; no crash	User stays on same page; no crash	Pass					

Post Condition: Essay is not evaluated. User receives **error feedback** without system crash, and app remains stable.

Test Case Essay ID: 19

Essay: In the passage, the author argues that public libraries remain a vital part of society despite the rise of digital technology. To support this claim, the author uses emotional appeals, facts, and expert opinions.

The author begins by appealing to the reader's sense of nostalgia, describing the library as "a safe place where learning meets community." This emotional connection makes the reader more likely to support the preservation of libraries. Additionally, the author strengthens the argument by providing statistical evidence citing a study showing that over 60% of people still use libraries regularly for free access to books, internet, and job resources.

To further build credibility, the author quotes respected educators and urban planners who emphasize the role libraries play in bridging the digital divide. These experts argue that not everyone can afford digital devices or fast internet, making public libraries essential.

Through a mix of emotional language, factual support, and credible expert voices, the author builds a convincing case for maintaining public libraries. The logical structure and balanced reasoning help persuade readers of the ongoing value of these institutions.

Table 6.19: Test case 19

Test Case ID:- 19			Test Design By: Aleena Sajid			
Test Module Name: System Integration			Test Design Date: 03/07/2025			
Test Priority:- High			Test Executed By: Aleena Sajid			
Test Title/Name: Valid IELTS Essay Submitted			Test Executed Date: 03/07/2025			
Description: Ensure essay goes from Node.js to Django, gets evaluated, and returns results successfully						
Pre-Condition: Essay and test type selected in Node.js						
Dependencies: Django API live and reachable						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Submit valid IELTS essay via UI	IELTS + Essay	Payload sent to Django	Payload sent to Django	Pass	
2	Django evaluates and returns JSON		Django responds in <15s	Django responds in <15s	Pass	
3	Node.js UI displays results		All fields shown correctly	All fields shown correctly	Pass	
Post Condition: The essay is scored and all IELTS fields (TR, CC, LR, GRA, Overall) + feedback appear in the UI.						

Table 6.20: Test case 20

Test Case ID:- 20			Test Design By: Aleena Sajid			
Test Module Name: System Integration			Test Design Date: 03/07/2025			
Test Priority:- Medium			Test Executed By: Aleena Sajid			
Test Title/Name: Missing Essay Field in Submission			Test Executed Date: 03/07/2025			
Description: Check how system handles a request with missing or null essay field						
Pre-Condition: User attempts submission with blank field						
Dependencies: Input validation in Node.js and Django						
Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Submit with empty essay	essay: ""	Validation triggered	Validation triggered	Pass	
2	API call blocked or returns 400		Error: "Essay field is required."	Error: "Essay field is required."	Pass	
3	Display error in UI		User sees red alert message	User sees red alert message	Pass	
Post Condition: Essay is not sent to backend, and the user receives a clear validation error.						

Table 6.21: Test case 21

Test Case ID:- 21	Test Design By: Aleena Sajid	
Test Module Name: System Integration	Test Design Date: 03/07/2025	
Test Priority:- Medium	Test Executed By: Aleena Sajid	
Test Title/Name: Invalid Test Type Sent to Backend	Test Executed Date: 03/07/2025	
Description: Ensure API rejects unknown test types (e.g., "TOEFL")		
Pre-Condition: UI sends unsupported test type		
Dependencies: Backend validation of test types		

Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Select invalid type manually or via dev tools	test_type: "TOEFL"	Invalid payload	Invalid payload	Pass	
2	Backend rejects request	400 Bad Request	Pass	Pass	Pass	
3	UI shows appropriate error		"Invalid test type selected"	"Invalid test type selected"	Pass	
Post Condition: No processing occurs. System validates test type strictly and blocks unregistered exams.						

Test Case Essay ID: 22

IELTS + Essay: In the passage, the author argues that public libraries remain a vital part of society despite the rise of digital technology. To support this claim, the author uses emotional appeals, facts, and expert opinions.

The author begins by appealing to the reader's sense of nostalgia, describing the library as "a safe place where learning meets community." This emotional connection makes the reader more likely to support the preservation of libraries. Additionally, the author strengthens the argument by providing statistical evidence citing a study showing that over 60% of people still use libraries regularly for free access to books, internet, and job resources.

To further build credibility, the author quotes respected educators and urban planners who emphasize the role libraries play in bridging the digital divide. These experts argue that not everyone can afford digital devices or fast internet, making public libraries essential.

Through a mix of emotional language, factual support, and credible expert voices, the author builds a convincing case for maintaining public libraries. The logical structure and balanced reasoning help persuade readers of the ongoing value of these institutions.

Table 6.22: Test case 22

Test Case ID:- 22	Test Design By: Aleena Sajid
Test Module Name: System Integration	Test Design Date: 03/07/2025
Test Priority:- High	Test Executed By: Aleena Sajid
Test Title/Name: Handle API Timeout or Network Error	Test Executed Date: 03/07/2025
Description: System response when Django is unreachable	
Pre-Condition: Network/API is down or misconfigured	
Dependencies:	

Step	Test Step	Test Data	Expected Result	Actual Result	Status (Pass/Fail)	Notes
1	Submit essay from Node.js	IELTS + Essay	Request fails	Request fails	Pass	
2	Axios/fetch throws timeout/error		Frontend handles gracefully	Frontend handles gracefully	Pass	
3	Show fallback message		“Unable to connect to evaluation service.”	“Unable to connect to evaluation service.”	Pass	
Post Condition: User is notified of connectivity issues. No crash, no retry loop. UX remains smooth.						

6.2 Decision Tree

Table 6.23: Decision Tree

Conditions	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Rule 8
First 500 Customers (20%)	T	F	T	F	F	F	T	F
1-year subscription (10%)	T	T	F	F	T	T	T	F
Coupons (15%)	T	T	T	T	T	F	F	F
Actions								
Discount (%)	30	X	20	X	15	20	30	X

6.2.1 Code Snippet

```
submitBtn.addEventListener("click", async () => { // on submit button click
  clearInterval(issueInterval); // stop issue essay timer
  clearInterval(argumentInterval); // stop argument essay timer

  submitBtn.disabled = true; // disable the submit button
  submitBtn.classList.remove("enabled"); // remove styling for enabled button

  const issue = issueTextarea.value.trim(); // get issue essay text
  const arg = argumentTextarea.value.trim(); // get argument essay text

  const prompt = `Prompt 1:\n${greIssuePromptText}\n\nPrompt 2:\n${greArgumentPromptText}`.trim(); // combine prompts
  const essay = `Essay 1:\n${issue}\n\nEssay 2:\n${arg}`.trim(); // combine both essays

  resultBox.style.display = "block"; // show result box
  window.scrollTo(0, 0); // scroll to top
  resultOutput.textContent = "Evaluating... Please wait."; // show loading message

  try {
    const res = await fetch("/send-essay", { // send essay to backend API
      method: "POST",
      headers: {"Content-Type": "application/json"}, // set content type
      body: JSON.stringify({ essay, test_type: "GRE", prompt }) // send essay, test type, and prompt
    });

    let out = (await res.json()).evaluation; // get evaluation result

    let clean = out.match(/Clarity:.*?Improvement:\s*(?:-|\*) .*(?=\n{2,}|$)/s); // extract key feedback section
    clean = (clean ? clean[0] : out) // use cleaned or full output
      .replace(/\[\*\]\s*/gm, '• ') // format bullets
      .replace(/^(Clarity|Logic|Development|Grammar|Total Score|Feedback.*?):/gm, '<strong>$1:</strong>'); // bold section headings

    resultOutput.innerHTML = clean; // show formatted result
  } catch {
    resultOutput.textContent = "Error during evaluation."; // error fallback
  }
});
```

Fig 6.1 Sending GRE Essays to Backend for AI Scoring

```

submitBtn.addEventListener("click", async () => { // on submit
  clearInterval(timer); // stop countdown
  submitBtn.disabled = true;
  submitBtn.classList.remove("enabled");

  const essay = textarea.value.trim(); // get essay text
  resultBox.style.display = "block"; // show result container
  window.scrollTo(0, 0);
  resultOutput.textContent = "Evaluating... Please wait.";

  try {
    const response = await fetch("/send-essay", { // send essay to backend
      method: "POST",
      headers: { "Content-Type": "application/json" },
      body: JSON.stringify({
        essay,
        test_type: "IELTS",
        prompt: ieltsPromptText // send prompt too
      })
    });

    const data = await response.json();
    let output = data.evaluation;

    let cleaned = output.match(/Task Response:.*?Improvement:\s*(?:-|\s)*.*?(?=\n{2,})$/s); // extract feedback
    cleaned = (cleaned ? cleaned[0] : output)
      .replace(/^[*\-]\s*/gm, '• ') // format bullets
      .replace(/^(Task Response|Coherence and Cohesion|Lexical Resource|Grammar|Overall Band Score|Feedback.*?):/gm, '<strong>$1:</strong>'); // bold headings

    resultOutput.innerHTML = cleaned; // show formatted output
  } catch {
    resultOutput.textContent = "An error occurred while evaluating the essay.";
  }
});

```

Fig 6.2 Logic for Submitting and Evaluating IELTS Essay Using AI

```

re.js > ...
submitBtn.addEventListener("click", async () => { // on submit click
  clearInterval(timer); // stop countdown timer
  submitBtn.disabled = true;
  submitBtn.classList.remove("enabled");

  const essay = textarea.value.trim(); // get typed essay
  resultBox.style.display = "block"; // show result box
  window.scrollTo(0, 0);
  resultOutput.textContent = "Evaluating... Please wait.";

  try {
    const response = await fetch("/send-essay", { // send to backend
      method: "POST",
      headers: { "Content-Type": "application/json" },
      body: JSON.stringify({ essay, test_type: "SAT", prompt: satPromptText })
    });

    const data = await response.json();
    let output = data.evaluation;

    let cleaned = output.match(/Thesis:.*?Improvement:\s*(?:-|\s)*.*?(?=\n{2,})$/s); // extract key block
    cleaned = (cleaned ? cleaned[0] : output)
      .replace(/^[*\-]\s*/gm, '• ') // format bullets
      .replace(/^(Thesis|Support|Organization|Evidence|Language Use|Total Score|Feedback.*?):/gm, '<strong>$1:</strong>'); // bold labels

    resultOutput.innerHTML = cleaned; // show formatted output
  } catch {
    resultOutput.textContent = "An error occurred while evaluating the essay.";
  }
});

```

Fig 6.3 Logic for Submitting and Evaluating SAT Essay Using AI

```

from django.urls import path
from .views import receive_essay, generate_prompt

urlpatterns = [
    path('receive_essay', receive_essay, name='receive_essay'),
    path('generate_prompt', generate_prompt, name='generate_prompt'),
]

```

Fig 6.4 API Route Configuration in urls.py

```

# views.py - Main Essay Evaluation Logic

from django.http import JsonResponse
from django.views.decorators.csrf import csrf_exempt
from django.views.decorators.http import require_http_methods
import requests, json

@csrf_exempt
@require_http_methods(["POST"])
def receive_essay(request):
    # Parse incoming JSON data
    data = json.loads(request.body)
    essay = data.get('essay', '')
    test = data.get('test_type', '').upper()
    prompt = data.get('prompt', '')

    # Basic input validation
    if not essay or test not in ['IELTS', 'SAT', 'GRE']:
        return JsonResponse({'error': 'Invalid input'}, status=400)

    # Set scoring instructions per test type
    instructions = {
        'IELTS': f"You are an IELTS examiner...\n{prompt}",
        'SAT': f"You are an SAT scorer...\n{prompt}",
        'GRE': f"You are grading GRE essays...\n{prompt}"
    }

    # Prepare payload for LLaMA2 API
    payload = {
        "model": "llama2",
        "messages": [
            {"role": "system", "content": instructions[test]},
            {"role": "user", "content": essay}
        ],
        "stream": False
    }

    # Send request to LLaMA2 for evaluation
    res = requests.post("http://localhost:11434/api/chat", json=payload)

    # Extract and return the evaluation reply
    reply = res.json().get("message", {}).get("content", "").strip()
    return JsonResponse({'evaluation': reply}, status=200)

```

Fig 6.5 Simplified `views.py` function for LLaMA2-based essay evaluation

6.3 Traceability Matrix

6.3.1 RID vs UCID (requirements vs use cases)

Table 6.24: RID vs UCID

RID/UCID	UC1	UC2	UC3	UC4	UC5	UC6	UC7
FR1	✓	✓	✓				✓
FR2	✓		✓		✓		✓
FR3			✓	✓			
FR4	✓	✓	✓				
FR5	✓		✓				
FR6		✓					
FR7	✓			✓			
FR8	✓						✓
NFR1		✓	✓				
NFR2		✓	✓			✓	
NFR3						✓	
NFR4							✓

6.3.2 Prototypes (RID vs PID)

Table 6.25: RID Vs PID

RID/PID	PID1	PID2	PID3	PID4
FR1			✓	
FR2			✓	
FR3			✓	
FR4				✓
FR5				✓
FR6				✓
FR7				✓
FR8		✓		
NFR1			✓	
NFR2				✓
NFR3		✓		
NFR4	✓			

6.3.3 Test Cases (RID vs TID)

Table 6.26: RID vs TID

R I D / T I D	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20	T21	T22
FR1	✓			✓	✓	✓	✓	✓					✓	✓	✓				✓			✓
FR2				✓		✓	✓	✓					✓	✓	✓				✓			✓
FR3			✓																			
FR4	✓							✓							✓				✓			✓
FR																				✓		

5																							
F R 6	✓								✓						✓				✓				✓
F R 7													✓										
F R 8	✓		✓	✓					✓	✓	✓				✓	✓			✓				
N F R 1												✓								✓			
N F R 2																		✓	✓				✓
N F R 3																	✓						
N F R 4																✓					✓		

6.3.4 Coverage (UCID vs TID)

Table 6.27: UCID vs TID

UC ID/ TID	T 1	T 2	T 3	Y 4	T 5	T 6	T 7	T 8	T 9	T 10	T 11	T 12	T 13	T 14	T 15	T 16	T 17	T 18	T 19	T 20	T 21	T22
UC 1	✓	✓	✓					✓	✓	✓					✓				✓		✓	
UC 2		✓							✓											✓	✓	
UC 3	✓			✓	✓	✓	✓	✓					✓	✓	✓				✓			✓
UC 4																						
UC 5											✓	✓				✓	✓	✓		✓	✓	✓
UC 6									✓													
UC 7														✓								

7. Results/Output/Statistics

7.1 %completion

The Education Testing Service web application is focused on essay evaluation for IELTS, SAT, and GRE exams. The writing module is fully completed, including rubric-based evaluation, scoring, and feedback features for all three test types. However, the full system also includes components for reading, listening, and speaking, which are **not implemented yet**. Therefore, considering the overall scope, we estimate the project to be **25% completed**, with the writing module being **100% complete**.

7.2 %accuracy

The backend (admin side) is functioning correctly. All features related to essay evaluation are integrated properly:

- Essays are received through the Node.js frontend
- Sent to the Django backend for evaluation

- Scored accurately based on respective rubrics (IELTS, SAT, GRE)
- Feedback is generated and sent back to the frontend.

The admin panel is fully functional it allows essay management, database updates, and tracking. Some **minor adjustments are still in progress on the frontend client side** to improve the display and UI for users. Considering the completed backend and the partially pending frontend polish, we estimate an overall system **accuracy of 80%**.

7.3 %correctness

The system functions correctly for its core functionality: Essay evaluation.

- The IELTS essays are scored using five metrics (Task Response, Coherence & Cohesion, Lexical Resource, Grammatical Range & Accuracy, Overall Band).
- SAT and GRE evaluations follow their respective rubrics.
- The interaction between Node.js and Django works seamlessly, and the UI is clean, simple, and user-friendly.

Although the full testing system isn't complete (reading, speaking, listening modules are pending), the essay evaluation component is fully functional and tested.

Hence, we estimate the correctness of the completed portion (writing system) at **100%**, and the overall project correctness at **40%**.

8. Conclusion

The developed web application effectively automates the evaluation of essays for IELTS, SAT, and GRE exams using their respective standardized rubrics. Users can submit their essays through a Node.js frontend, which are then processed by a Django backend. The system generates accurate scores and detailed feedback covering aspects such as task response, coherence, vocabulary, grammar, and organization.

During the development process, we experimented with several machine learning and deep learning models to achieve the most accurate and consistent evaluation results. After extensive testing, we finalized our backend evaluation engine using the **LLaMA** model, which provided the best balance of accuracy, fluency, and contextual understanding. For SAT and GRE essays, we used the **DeepSeek API** to generate rubric-aligned evaluations and improvement suggestions. This hybrid approach ensured high-quality output while maintaining performance efficiency.

The system is user-friendly, scalable, and supports both users and admins. The writing module is **100% complete**, delivering robust performance across all supported exam types. Overall, the project demonstrates a strong implementation of AI-assisted essay evaluation and can serve as a valuable tool for students and educational institutions.

9.FUTURE WORK

While the writing module is fully implemented, the broader goal of building a complete educational testing platform requires the following future enhancements:

1) Reading Module

- Integrate reading passages with automatic question generation and evaluation.
- Implement score calculation based on comprehension accuracy.

2) Listening Module

- Support audio-based question delivery with automated response evaluation.
- Add timing and speech recognition functionalities.

3) Speaking Module

- Enable voice input for speaking tests.
- Use AI (e.g., Whisper or similar) for evaluating pronunciation, fluency, and coherence.
- Store recordings for admin review and manual scoring if needed.

4) Enhanced User Analytics

- Track user performance across different tests over time.
- Visualize insights and provide personalized improvement suggestions.

5) Model Optimization

- Continue refining the LLaMA-based evaluation engine for greater precision.
- Expand DeepSeek API integration and consider domain-specific fine-tuning.
- Explore multimodal input handling in future versions.

6) Mobile App Development

- Launch a mobile version of the platform to enhance accessibility.
- Ensure responsive design, offline support, and push notifications.

With these future enhancements, the application can evolve into a comprehensive AI-powered language testing system that evaluates all four core skills writing, reading, listening, and speaking in a seamless and scalable manner.

10. Bibliography

10.1 Books

None.

10.2 Journals

None.

10.3 Articles

- 1) Graduate Admission Prediction Using Machine Learning --> https://www.researchgate.net/profile/Ali-Nassif/publication/348433004_Graduate_Admission_Prediction_Using_Machine_Learning/links/5ffee71745851553a03dd015/Graduate-Admission-Prediction-Using-Machine-Learning.pdf.

10.4 Research papers

1. Comparing expert tutor evaluation of reflective essays with marking by generative artificial intelligence (AI) tool --> <https://www.sciencedirect.com/science/article/pii/S2666920X24000274>.
2. A Hybrid Automatic Scoring System: Artificial IntelligenceBased Evaluation of Physics Concept Comprehension Essay Test --> <https://www.ijiet.org/vol14/IJiet-V14N6-2113.pdf>.
3. Human-AI Collaborative Essay Scoring: A Dual-Process Framework with LLMs --> <https://arxiv.org/abs/2401.06431>.
4. US College & University Admissions 2020-2021 --> <https://www.kaggle.com/datasets/jfschultz/us-college-admissions-2021-rates-and-test-scores>.
5. Internationals Applying to US Colleges --> <https://www.kaggle.com/datasets/priyanshusethi/internationals-applying-to-us-colleges>.
6. ieomsociety Paper --> <https://ieomsociety.org/dc2018/papers/383.pdf>.

10.5 Other References

1. Automatic Essay Scoring --> <https://github.com/sankalp99/Automatic-Essay-Scoring>.
2. EssayScore_FYP --> https://github.com/tingwei3931/EssayScore_FYP/tree/master?tab=readme-ov-file.
3. DeepEssay --> <https://github.com/Logisx/DeepEssay>.
4. Automated-essay-grading --> <https://github.com/hazem-alabiad/automated-essay-grading-system/tree/master>.
5. Essay_Grading_IELTS --> <https://github.com/ADHIKSHA/Essay-Grading-IELTS/blob/master/requirements.txt>.
6. Transformer-Bert --> <https://github.com/krishnaik06/Trnasformer-Bert>.
7. Automatic-Essay-Grading-and-Feedback --> <https://github.com/aniruddhadave/Automatic-Essay-Grading-and-Feedback>.
8. English Language Learning Prediction with AI --> <https://github.com/suhasmaddali/English-Language-Learning-Prediction-with-AI-and-Machine-Learning>.
9. Graduate Admission Prediction --> <https://www.kaggle.com/datasets/mukeshmanral/graduates-admission-prediction>.

10. Graduate Admission 2 --> <https://www.kaggle.com/datasets/mohansacharya/graduate-admissions>.
11. Data from Edulix.com website --> <https://github.com/joemanley201/universityRecommendationSystem/tree/master/scrapper/univJSON>.
12. Graduate-Admission-Prediction --> <https://github.com/srinivasav22/Graduate-Admission-Prediction/tree/master>.
13. Chance-of-Graduate-Admission-Prediction --> <https://github.com/r24zeng/Chance-of-Graduate-Admission-prediction>.
14. Graduate-admission-prediction-using-ANNs --> <https://github.com/BhaveshtBhakta/Graduate-Admission-Prediction-Using-ANN>.
15. Graduate-Admission-Chances --> <https://github.com/dhanavikram/Graduate-Admission-Chances/tree/main>.
16. University-admission-prediction --> <https://github.com/shloak17107/University-Admissions-Prediction/tree/master>.

11. Appendix

11.1 Glossary of terms

Table 11.1: Glossary of terms

Term	Definition
IELTS	International English Language Testing System evaluates English proficiency in listening, reading, writing, and speaking.
SAT	Scholastic Assessment Test standardized test for college admissions in the U.S., includes a writing section.
GRE	Graduate Record Examination standardized test for graduate admissions, includes an analytical writing task.
LLaMA	Large Language Model Meta AI an open-source language model used for essay evaluation due to its accuracy and performance.
DeepSeek	A language model API used for SAT and GRE essay evaluation, providing scores and feedback based on rubrics.
Node.js	A JavaScript runtime environment used for building the frontend and handling user essay submissions.
Django	A Python-based backend framework used to process essays, evaluate them using AI models, and return results.
Rubric	A set of scoring guidelines used to evaluate the quality of writing in standardized tests.
Task Response (TR)	IELTS scoring criterion evaluating how well the essay addresses the task and presents relevant arguments.
Coherence and Cohesion (CC)	IELTS criterion assessing the logical organization and flow of ideas in the essay.
Lexical Resource (LR)	IELTS criterion that measures the range and accuracy of vocabulary used in the essay.
Grammatical Range & Accuracy (GRA)	IELTS criterion evaluating the use, variety, and correctness of grammar.

Analytical Writing	A GRE task requiring the analysis of an issue or argument in a structured essay format.
Frontend	The part of the application users interact with (built using Node.js) to submit essays and view evaluations.
Backend	The server-side of the application (built using Django) responsible for essay processing and interaction with AI models.
API	Application Programming Interface enables communication between systems such as frontend, backend, and AI services like DeepSeek.
Evaluation Engine	The core module that uses AI models (LLaMA, DeepSeek) to generate scores and feedback based on exam-specific rubrics.

11.2 Pre-requisites

To set up and run the essay evaluation system, the following development and deployment steps must be completed:

1. Clone the Project

- Download the full source code from the GitHub repository.

2. Set Up Python Environment (Backend)

- Install Python 3.9 or higher.
- Navigate to the Django backend directory.
- Create a virtual environment and activate it.
- Install required dependencies using:

```
pip install -r requirements.txt
```

3. Set Up Node.js Environment (Frontend)

- Install Node.js version 16 or higher.
- Navigate to the frontend directory.
- Install dependencies using:
 - `npm install`

4. Download and Set Up LLaMA Model (via Ollama)

- Visit the official [Ollama website](#) to download the LLaMA model.
- Follow Ollama's setup instructions to install the model locally.
- Start the LLaMA server using:


```
ollama run llama2
```

5. Run Django Server

- Navigate to the Django project directory and start the backend server:


```
python manage.py runserver
```

6. Run Node.js Server

- In the frontend directory, start the Node.js server to launch the user interface:


```
npm start
```

7. Request Handling Flow

- Users submit their essays through the Node.js frontend --> Essays are sent to the Django backend via HTTP requests --> Essays are sent to the Django backend via HTTP requests. --> The backend processes the essay using the LLaMA model (or DeepSeek API) and returns scores and feedback.