AI-POWERED STUDENT GRADING AND TEACHER RATING AUTOMATION SYSTEM

25-26J-162

Project Proposal Report

Jayasooriya L.T

BSc. (Hons) Degree in Information Technology, Specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

August 2025

AI-POWERED STUDENT GRADING AND TEACHER RATING AUTOMATION SYSTEM

25-26J-162

Project Proposal Report

Jayasooriya L.T – IT22095480

Supervised By Prof. Samantha Rajapaksha

Co-Supervised by Mrs. Bhagyani Chathurika

BSc. (Hons) Degree in Information Technology, Specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

August 2025

I. Declaration

We declare that this is our own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

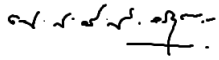
Name	Student ID	Signature
Jayasooriya L. T	IT22095480	

Supervisor's statement: The candidates mentioned above are conducting research for their undergraduate dissertations under my supervision.

Supervisor

(Prof. Samantha Rajapaksha)

Date: 8/26/2025



Co-supervisor

(Mrs. Bhagyani Chathurika)

Date: 8/26/2025

II. Abstract

This research proposes the development of an AI-powered system that automates student grading and provides intelligent teacher reteaching guidance using custom-built Optical Character Recognition (OCR), Natural Language Processing (NLP), and Machine Learning (ML) models. Current educational assessment methods are often time-consuming, subjective, and inconsistent. Moreover, teachers lack structured guidance on which concepts to reteach and in what sequence after identifying student weaknesses. The proposed system tackles these problems by integrating a web and mobile application for document capture, automated evaluation, and AI-driven reteaching recommendations.

The solution starts with an OCR module that extracts text from scanned or photographed answer scripts and feedback forms. It supports both handwritten and printed content. The AI grading engine processes the extracted responses, applying NLP and ML techniques to assess answers based on marking schemes, semantic similarity, structure, and language quality. The system generates automated scores, question-wise explanations, and personalized improvement recommendations for students.

For teacher reteaching guidance, the platform utilizes a Knowledge Graph-based approach that maps student errors to syllabus concepts and prerequisite relationships. Using Graph Neural Networks (GNNs) and Bayesian Knowledge Tracing (BKT), the system propagates weaknesses across concept dependencies and predicts future learning risks. The AI-powered reteaching module generates targeted recommendations for teachers, identifying which prerequisite concepts need reinforcement and providing visual heatmaps overlaid on the Knowledge Graph to guide instructional decisions at both individual student and classroom levels.

The methodology includes developing a custom OCR pipeline, creating NLP-based grading models, constructing syllabus-based Knowledge Graphs, and implementing GNN-BKT hybrid models for weakness propagation analysis. The expected results include significant reductions in grading time, better feedback accuracy, and data-driven reteaching strategies that ensure students master foundational concepts before advancing. The system's offline capabilities and mobile support ensure it works well in low-connectivity environments. This research aims to deliver a scalable, objective, and efficient academic evaluation platform that benefits both students and educators through intelligent automation and structured pedagogical guidance.

Keywords: AI-powered grading, OCR, NLP, teacher performance evaluation, educational analytics

Table of Contents

List of figures

List of tables

List of abbreviations

Abbreviation	**Full Form**
AI	Artificial Intelligence
OCR	Optical Character Recognition
NLP	Natural Language Processing
ML	Machine Learning
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
API	Application Programming Interface
SUS	System Usability Scale
DL	Deep Learning
DB	Database
UI	User Interface
UX	User Experience

1. Introduction

In the changing educational landscape, timely evaluation of both students and teachers is crucial for improving learning outcomes and maintaining teaching quality. Traditional assessment methods, like manual grading of answer sheets and teacher performance reviews, are often labor-intensive, subjective, and inconsistent. These drawbacks lead to delayed feedback and make it hard to provide specific, data-based improvements.

As online and remote learning environments grow, the need for objective, scalable, technology-based assessment systems has become more obvious. However, many current solutions focus only on grading students or offer limited evaluations for teachers, often relying on external APIs or internet access. These issues show the potential to use AI and Machine Learning techniques to make the evaluation process smoother and promote fairness and transparency.

This research presents an AI-driven system that automates both student grading and teacher performance evaluations. The platform includes a custom Optical Character Recognition (OCR) module for extracting content from printed and handwritten documents, Natural Language Processing (NLP) for semantic analysis, and Machine Learning algorithms for performance assessment. Students receive not just scores but also detailed explanations for each question along with personalized improvement recommendations. Teachers benefit from a multi-faceted evaluation method that looks at student performance trends, sentiment analysis of feedback, assessments based on Bloom's Taxonomy, and engagement metrics.

The system provides results through an integrated web and mobile application, offering offline functionality to tackle accessibility issues in low-connectivity areas. Overall, the research aims to create a fair, efficient, and data-driven evaluation framework that benefits educational institutions, teachers, and students.

1.1.Background & Literature survey

Assessment of students and evaluation of teacher performance are central to maintaining quality in education. However, conventional grading approaches are slow, require significant manual effort, and often vary between evaluators, which can reduce fairness and consistency [1], [2]. Similarly, teacher evaluations are frequently based on a narrow set of criteria, which can lead to biased or incomplete feedback that offers little direction for professional development [4]. These limitations become even more significant in large academic settings or in remote learning, where timely and scalable solutions are essential [1].

The integration of Artificial Intelligence (AI) and Machine Learning (ML) into educational systems has shown promising potential for resolving these challenges [1], [5]. One critical

enabling technology is Optical Character Recognition (OCR), which converts printed or handwritten content into digital text for automated processing [2]. While OCR is widely used, standard tools often fail to achieve high accuracy on handwritten academic scripts, especially those with varied layouts or low-quality scans. This shortcoming highlights the need for customized OCR pipelines supported by robust image preprocessing [2].

Natural Language Processing (NLP) methods have been applied to automatically assess descriptive responses. Techniques such as Named Entity Recognition (NER), semantic matching, and deep learning-based embeddings like BERT have been effective in evaluating answer quality against model solutions [1], [3]. Despite these advancements, most existing solutions provide only overall scores, lacking detailed, question-specific feedback and personalized learning pathways [1], [3], [5].

For teacher performance assessment, AI-driven approaches have introduced sentiment analysis of student feedback and analysis of performance trends [4]. Yet, many existing systems are limited by the absence of mobile access, offline functionality, or comprehensive evaluation models that combine multiple performance indicators. Incorporating frameworks such as Bloom's Taxonomy and response pattern detection could provide deeper, data-driven insights into teaching quality [4].

The research presented in this proposal seeks to fill these gaps by designing a self-contained, cross-platform academic evaluation system that combines custom OCR, NLP, and ML models. The system will deliver automated grading with question-wise explanations for students and multi-factor ratings for teachers, all within a platform capable of functioning offline.

1.2.Research Gap

Existing research on evaluating second language learning in Sri Lanka, such as "A Needs Analysis in Second Language Teaching (English Language): A Study on Grade 10,11 Students in the Sri Jayawardenapura Educational Zone,"[9] relies heavily on manual questionnaires, teacher observations, and periodic surveys. While these methods provide useful insights, they do not offer the scalability, objectivity, and real-time feedback that students need for continuous improvement.

These traditional methods often provide general feedback instead of personalized recommendations. This makes it harder for students to identify specific areas they struggle with in speaking, reading, writing, or grammar. Moreover, manual assessment is time-consuming and often biased, leading to delayed and inconsistent feedback.

There is also a lack of integrated, AI-based systems that analyze student performance continuously and deliver immediate, actionable suggestions. Current studies do not use machine learning to identify mistakes, detect misunderstandings, or provide real-time progress tracking.Additionally, most research doesn't include question-level evaluation, cross-skill

performance mapping, or automated analytics, all of which are important for effective learning support.

The proposed component aims to fill these gaps by introducing a machine learning-based performance analysis system. This system will continuously evaluate Grade 11 English Literature exam results, pinpoint learning deficiencies in detail, and create personalized improvement recommendations. Unlike previous approaches, this solution combines OCR-based exam data collection, real-time analytics, and AI-generated suggestions in one platform. This setup provides students and teachers with immediate access to useful insights. This innovation moves away from manual, survey-based evaluation towards a scalable, data-driven approach that enhances learning outcomes.

Feature	Proposed Research	Needs Analysis Study (Existing Research)
Analysis Method	AI-powered pattern recognition with ML	Manual questionnaires & teacher observations
Data Collection	Continuous automated data from exams via OCR	One-time survey (60 students)
Skill Gap Detection	AI-powered grading with question-wise explanations and personalized improvement suggestions	Manual teacher evaluation & self-report
Personalization	Automated, real-time detection	Generic recommendations
Coverage Scope	All English skills (reading, writing, speaking, listening, grammar)	Limited to speaking & listening
Feedback Delivery	Instant AI-generated suggestions via app/web	Delayed, teacher-dependent feedback
Progress Tracking	Continuous ML-based monitoring	Periodic manual evaluation

Table 1. Comparison of Proposed Research with Existing Studies

1.3.Research Problem

In educational environments, the evaluation of students and teachers is a critical process that influences learning outcomes, teaching quality, and institutional performance. However, existing systems face several key limitations.

For student evaluation, manual grading of answer scripts is time-consuming, inconsistent, and prone to human bias [1], [2]. Many available automated systems focus only on typed responses [6], cannot process handwritten answer sheets [2], and fail to deliver question-specific explanations or targeted recommendations for improvement [1], [3], [5]. As a result, students often receive delayed, generic feedback that does not adequately address their unique learning needs [1], [3].

For teacher evaluation, most platforms rely on limited performance metrics gathered through surveys or manual entries [7], [8]. They rarely incorporate advanced analytical approaches such as sentiment analysis of handwritten feedback [4], Bloom's Taxonomy-based question paper evaluation, or detection of student response patterns [4], [8]. Additionally, these systems often operate in isolation from student assessment modules [6], [7], preventing a unified view of academic performance.

The situation is further complicated by technical limitations in current solutions. Many depend heavily on cloud-based APIs or online services [6], [7], making them unsuitable for low-connectivity environments, and few offer mobile accessibility with offline functionality [6], [8].

Given these challenges, there is a clear need for a comprehensive, AI-powered evaluation platform that can:

1. Process both printed and handwritten answer scripts using a custom OCR pipeline [2].
2. Automate grading with question-level explanations and personalized improvement recommendations for students [1], [3], [5].
3. Evaluate teachers using a multi-factor model combining performance trends, sentiment analysis, Bloom's Taxonomy classification, and engagement metrics [4], [7], [8].
4. Function across web and mobile platforms with full offline capabilities to ensure accessibility in diverse learning environments [6], [8].

Addressing this problem will result in a fairer, faster, and more data-driven academic evaluation process that benefits students, teachers, and institutions alike.

2. Objectives

The primary objective of this research is to design and develop an AI-powered academic evaluation platform capable of automating student grading and teacher performance evaluation through custom-built Optical Character Recognition (OCR), Natural Language Processing (NLP), and Machine Learning (ML) models. This platform will address the limitations of traditional academic evaluation methods by integrating offline processing, detailed feedback mechanisms, and multi-factor analytics to ensure fair, efficient, and data-driven decision-making in education.

2.1.Main Objectives

Develop an integrated web and mobile application to process printed and handwritten academic scripts, generate question-wise grading with improvement suggestions for students, and provide multi-factor performance evaluations for teachers.

- Specific: The system will use OCR, NLP, and ML to process answer sheets, grade responses, give question-specific feedback, and evaluate teachers based on multiple metrics, including student performance trends, sentiment analysis, Bloom's Taxonomy classification, and engagement patterns.
- Measurable: System effectiveness will be measured by achieving at least 90% OCR accuracy, 85% grading accuracy (compared to human marking), and a 10% improvement in feedback turnaround time during pilot testing [1]– [8].
- Achievable: The objective is feasible with current open-source OCR tools (e.g., Tesseract), ML frameworks (e.g., TensorFlow, PyTorch), and standard development technologies for web/mobile platforms [1], [2].
- Realistic: The development plan aligns with available data, technical expertise, and infrastructure, making it achievable within the project's scope and timeline [4], [6].

Time-bound: The full system will be completed in 12 months, with 3 months for OCR and grading model development, 4 months for teacher evaluation model and mobile app, 3 months for integration and testing, and 2 months for refinements.

2.2.Specific Objective

1. Identify Student Improvement Areas through Performance Analysis

- Specific: Use ML models to analyze student performance data and detect weak learning areas, generating personalized improvement suggestions [1], [3].

- Measurable: Deliver 90% accuracy in identifying key improvement areas and at least 15% improvement in student scores during pilot testing [5].
- Achievable: Feasible using clustering, pattern recognition, and keyword matching techniques [3].
- Realistic: Builds upon existing AI-driven analytics frameworks that have been validated in educational contexts [6].
- Time-bound: Complete in 2 months, with 1 month for algorithm development and 1 month for testing with real student data.

By achieving these SMART objectives, the research will address critical gaps in academic evaluation systems by delivering a unified, offline-capable, AI-driven platform that provides fair, accurate, and actionable feedback for both students and teachers.

3. Methodology

The methodology of this research uses a structured, multi-phase approach to design and develop the AI-Powered Student Grading and Teacher Rating Automation System. The project combines OCR-based student grading and multi-factor teacher evaluation into a single platform. The process includes data collection, model development, system implementation, and evaluation to create a solid and accurate solution.

3.1.Research Design

This study uses a Design Science Research Methodology (DSRM) framework, focusing on developing and evaluating artifacts. The final artifact will be a web and mobile application that works on multiple platforms. It will include these components:

1. OCR module and student grading engine.
2. Teacher rating engine.
3. Student improvement analysis.
4. Teacher improvement analysis.

3.2.Individual System Diagram

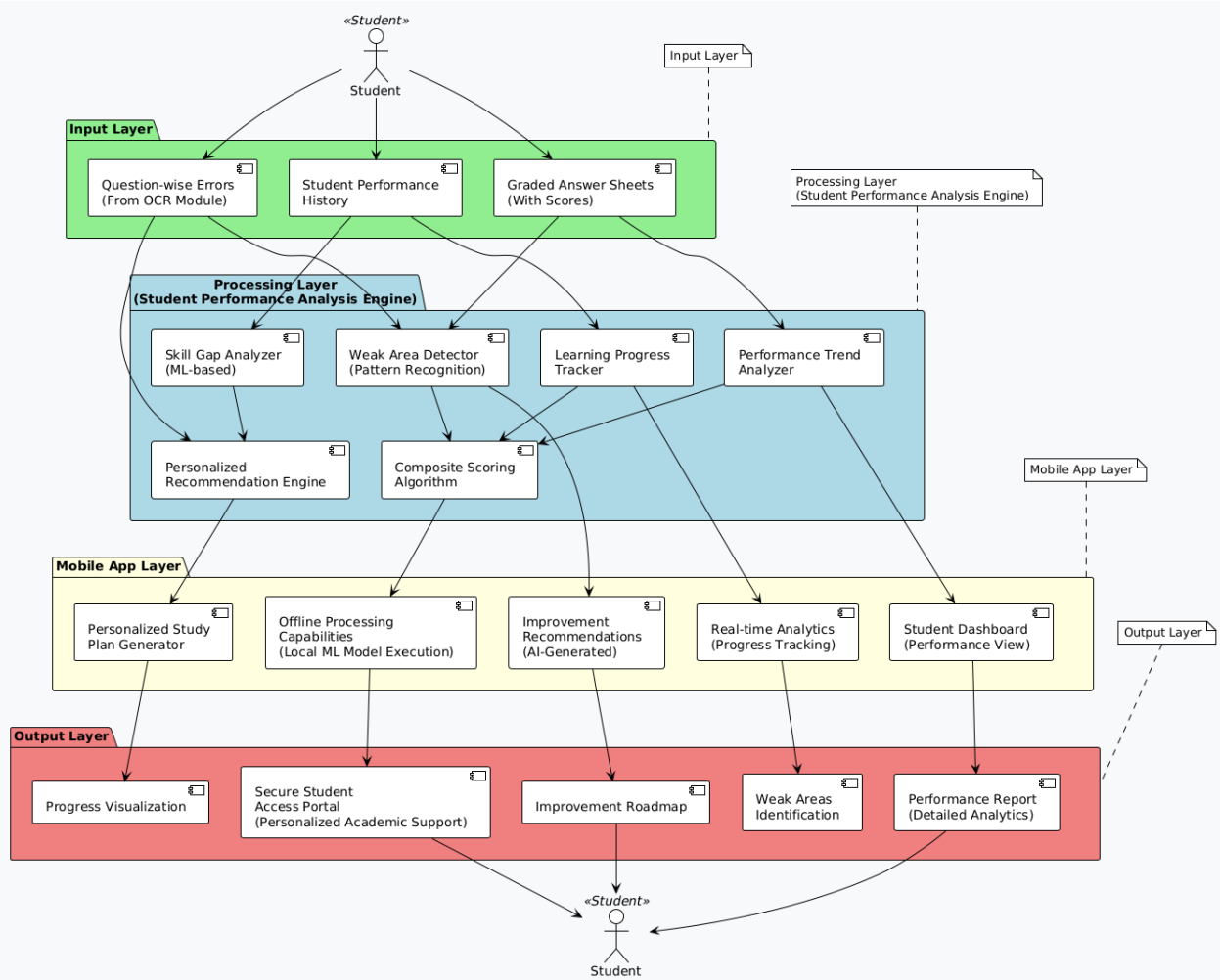


Figure 1. Individual System Diagram

3.3.Data Collection

The system will use ethically sourced and anonymized educational datasets.

These include:

- Handwritten and printed student answer sheets.
- Marking schemes and model answers.
- Teacher evaluation feedback forms.
- Historical student performance data.
- Tagged datasets for Bloom's Taxonomy classification.

We will gather datasets from partner institutions and manually label them for training and validation [1], [2].

3.4.System Architecture

The AI-Powered Student Grading and Teacher Rating Automation System will follow a five-layer architecture [6]-[8]:

1. Input Layer: Upload answer sheets and feedback forms (mobile/web).
2. Preprocessing Layer: Image preprocessing, including noise reduction and skew correction, using OpenCV [2].
3. Processing Layer: OCR, NLP-based grading, and ML-driven analytics [1], [3].
4. Analysis Layer:
    - Automated student grading with explanations.
    - Teacher evaluation using sentiment, Bloom's Taxonomy, and performance metrics [4], [7].
5. Output Layer: Personalized improvement reports, dashboards, and performance analytics.

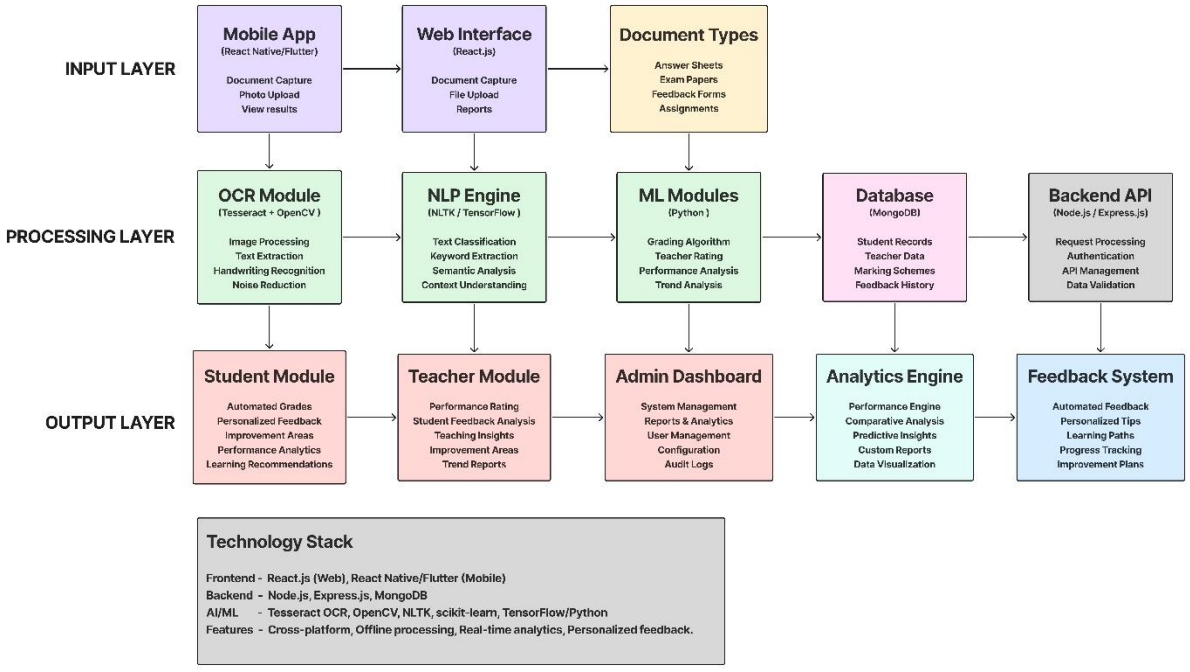


Figure 2. Overall System Architecture Diagram

3.5.Model Development

3.5.1. OCR Module & Student Grading Engine

- Develop a custom OCR pipeline to extract handwritten and printed text [2].
- Use NLP and ML models for semantic-based grading [1], [3].
- Provide scores and explanations for each question [5].

3.5.2. Student Improvement Analysis

- Identify weak learning areas using ML clustering [5].
- Provide personalized recommendations through mobile or web [6].
- Track ongoing improvement over time.
-

3.5.3. Teacher Improvement Analysis

- Use Bloom's Taxonomy and pattern detection to identify teaching gaps [8].
- Suggest specific improvements in teaching methods [4], [7].

3.6.System Implementation

The AI-Powered Student Grading and Teacher Rating Automation System will use a modular approach. This will improve scalability and maintainability. The backend will be built with Node.js and Express.js. This setup will manage APIs, integrate machine learning models, and handle communication between the database and the frontend. The frontend will utilize React.js for the web application and React Native for the mobile application. This ensures a responsive and cross-platform user experience.

For data storage, MongoDB will be the main database. It will hold answer sheets, grading results, teacher ratings, and performance records. To support users in low-connectivity areas, the system will be an offline first application. It will have synchronization tools to update data once an internet connection is available. This design provides reliability, speed, and ease of use for both students and teachers.

3.7.Evaluation Metrics

The effectiveness of the system will be measured using clear performance metrics. OCR accuracy will be evaluated to ensure at least 90% recognition accuracy for both handwritten and

printed text [2]. Grading accuracy will be validated by comparing the automated results with human evaluator scores, aiming for an 85% or higher correlation [1], [3]. Similarly, the teacher evaluation accuracy will be compared to expert assessments, targeting at least 85% agreement [4], [7].

Additionally, the system's usability will be assessed through the System Usability Scale (SUS) to measure user satisfaction and ease of use among teachers and students. The system's performance will also be tested to maintain an average processing time of less than 3 seconds per grading or evaluation request, ensuring efficiency in real-time scenarios.

3.8.Validation Strategy

To check how well the system works, we will use several testing methods. We will conduct pilot testing with Grade 11 English Literature answer sheets. This will let us evaluate OCR, grading, and the detection of improvements in a real-world setting. We will also carry out a comparative study against existing systems like EnglishBuddy [6], Automated Teacher Performance

Evaluation [7], and ML-Based Teaching Quality Assessment Systems [8]. This will help us point out improvements and new features.

In addition, we will gather expert reviews from experienced teachers and examiners. Their insights will help us confirm grading accuracy, feedback quality, and teacher evaluation results. This mixed validation approach makes sure the system is accurate, fair, and suitable for educational settings.

3.9.Time Frame

The project will be completed in 12 months. It has specific milestones, including 50% and 80% completion targets, which are detailed in the Gantt chart. Each phase will have its own time frame. This approach will help the project move forward smoothly and achieve its goals on schedule.

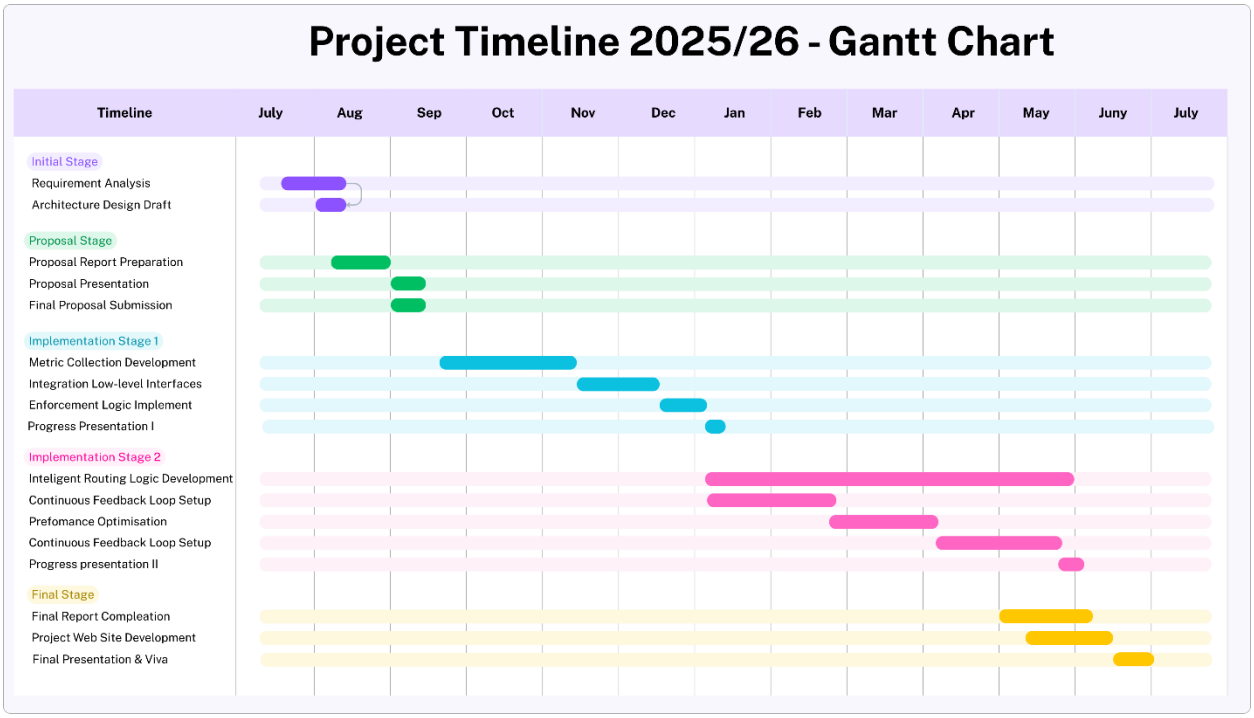


Figure 3. Project Grant Chart

4. Description of Personal and Facilities

4.1.Personnel

This project, AI-Powered Student Grading and Teacher Rating Automation System, is carried out by a team of four undergraduate researchers specializing in Information Technology at SLIIT. Each member is responsible for a key part of the system:

- **Pathiraja P.U.M (IT22243362):** Development of the OCR module and student grading engine, focusing on text extraction from handwritten and printed academic scripts.
- **Wanniarachchi W.A.P.M (IT22103154):** Implementation of the Teacher Reteaching automation engine, integrating with AI-driven evaluation models like GKT, GNN with Knowledge Graps(KG).
- **Jayasooriya L.T (IT22095480):** Design of algorithms for student performance analysis, identifying weak areas and generating personalized improvement recommendations.
- **Hettiarachchi R.H (IT22120052):** Development of teacher improvement analytics using Bloom's Taxonomy, sentiment analysis, and performance monitoring.

The research team is supervised by Prof. Samantha Rajapaksha, who offers guidance on research methods, technical validation, and academic standards.

Personnel – Individual Component (Jayasooriya L.T, IT22095480)

Jayasooriya L.T is mainly responsible for designing and developing the **Student Performance Analysis module** in the *AI-Powered Student Grading and Teacher Rating Automation System*. His work focuses on building algorithms to analyze student responses, detect weak areas, and generate **personalized improvement recommendations** based on performance trends.

The module leverages **Machine Learning (ML)** techniques to identify recurring mistakes and patterns in student answers, while also integrating **Bloom's Taxonomy** to classify responses according to cognitive skill levels. This allows the system to go beyond traditional grading by offering a deeper understanding of each student's strengths and weaknesses.

In addition, he will develop mechanisms to track **student progress over time**, presenting detailed reports and visualizations that help both students and teachers monitor improvement. The system will also incorporate **adaptive learning feedback**, ensuring that recommendations are tailored to the individual learning path of each student.

Through this contribution, Jayasooriya L.T ensures that the platform not only automates grading but also supports **continuous learning and personalized guidance**, creating a more effective and data-driven academic evaluation system.

.

4.2.Facilities

The successful implementation of the AI-Powered Student Grading and Teacher Rating Automation System needs a mix of software tools, hardware resources, and school infrastructure. The project will rely on modern software, using frameworks like Node.js, Express.js, React.js, and React Native/Flutter for backend and frontend development. For the artificial intelligence parts, TensorFlow, PyTorch, and Scikit-learn will design and train machine learning models. Tesseract and OpenCV will support the custom OCR module. A MongoDB database will act as the primary data storage solution. Collaboration platforms like GitHub, Slack, and Google Drive will be used for version control, communication, and document management.

For hardware resources, the project team will use high-performance computers with GPU acceleration to meet the computational needs of training and testing deep learning models. Mobile devices on both Android and iOS will be used for testing the cross-platform mobile application. For preparing datasets, scanners and digital cameras will capture and digitize handwritten answer sheets, creating a varied training and validation dataset.

The project will also gain strong support from the institution with access to laboratories, computing infrastructure, and technical resources available at SLIIT. These facilities will create a secure and collaborative environment for developing, testing, and integrating models. The institution will also offer expert help, especially for labeling datasets, validating grades, and evaluating teacher models. Additionally, secure storage will ensure that all student and teacher data stay anonymous and protected, following ethical and institutional data-handling standards.

5. Budget and Budget Justification

Description	Cost	Occurring
Server Hosting (VPS/Cloud)	LKR 8,000	Monthly
App Store Hosting (Apple)	LKR 7,754	One-Time
Play Store Hosting (Google)	LKR 30,700	Annual
Cloud Database (Amazon S3)	USD 0.023 per GB	Monthly
GPU Cloud Service (Model Training)	LKR 15,000	Monthly (as required)
Domain Registration & SSL	LKR 5,000	Annual
Miscellaneous (Testing devices, Scanning tools, Internet, etc.)	LKR 5,000	One-Time

Table 2. Budget For the Project

Budget Justification

The proposed budget outlines the necessary costs for implementing and deploying the AI-Powered Student Grading and Teacher Rating Automation System. A monthly server hosting fee of LKR 8,000 is needed to maintain the backend and ensure system reliability. The Apple App Store hosting fee of LKR 7,754 is required as a one-time expense, while the Google Play Store fee of LKR 30,700 is an annual cost. Both are crucial for distributing mobile applications to users. For secure and scalable data storage, Amazon S3 will be used, with costs estimated at USD 0.023 per GB per month, depending on the size of the dataset.

Since the project involves training machine learning models, we may need additional computing power from GPU-based cloud services, estimated at LKR 15,000 per month, based on usage. Domain registration and an SSL certificate cost LKR 5,000 annually and provide secure access to the system. Finally, we have included miscellaneous costs of LKR 12,000 as a one-time fee to cover expenses for testing devices, document scanning tools, and internet access during the development and validation phases.

This budget ensures that the system can be developed, deployed, and maintained effectively while meeting the project's technical and functional requirements

6. Reference

[1] K. Rani and R. Sujatha, “Automated Evaluation of Descriptive Answers using NLP Techniques,” *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 5, pp. 435–440, 2020.

[2] M. Jain and V. Yadav, “Optical Character Recognition Using Tesseract and OpenCV,” *Journal of Emerging Technologies and Innovative Research*, vol. 6, no. 2, pp. 112–116, 2019.

[3] S. Poria, E. Cambria, and A. Gelbukh, “Aspect Extraction for Opinion Mining with a Deep Convolutional Neural Network,” *Knowledge-Based Systems*, vol. 108, pp. 42–49, 2016.

[4] M. Balaji and R. Lavanya, “Machine Learning-Based Feedback Analysis for Teacher Evaluation,” *Procedia Computer Science*, vol. 189, pp. 467–474, 2021.

[5] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[6] Gurusinghe et al., “EnglishBuddy – Automated English Language Learning Platform,” *International Research Journal*, ISSN 2659-2193, 2019.

[7] IEEE, “Automated Teacher Performance Evaluation Using ML Algorithms,” *IEEE Access*, 2021.

[8] IEEE, “ML-Based Teaching Quality Assessment System,” *IEEE Transactions on Learning Technologies*, 2022.

[9] T. Boteju and U. Rathnasena, “A Needs Analysis in Second Language Teaching (English Language): A Study on Grade 10,11 Students in the Sri Jayawardenapura Educational Zone,” in *Proc. 4th SLIIT Int. Conf. Advancements in Sciences and Humanities (SICASH)*, Colombo, Sri Lanka, Dec. 2023, pp. 91–96.