

***Dissertation on***

**“Capstone Tracker with an Integrated Evaluation System”**

*Submitted in partial fulfillment of the requirements for the award of the degree*

*of*

**Bachelor of Technology**

**in**

**Computer Science & Engineering**

**UE22CS320B – Capstone Project Phase - 2**

***Submitted by:***

**Dhanush S Jettipalle PES2UG22CS175**

**Ketan Kancharla PES2UG22CS263**

**Nitheesh Pugazhanthi PES2UG22CS371**

**Rohan M G PES2UG22CS454**

*Under the guidance of*

|  |
| --- |
| **Dr.Richa Sharma** Associate Professor PES University |

**January - May 2025**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

FACULTY OF ENGINEERING

**PES UNIVERSITY**

(Established under Karnataka Act No. 16 of 2013)

Electronic City, Hosur Road, Bengaluru – 560 100, Karnataka, India

Logo, company name

Description automatically generated

**PES UNIVERSITY**

(Established under Karnataka Act No. 16 of 2013)

Electronic City, Hosur Road, Bengaluru – 560 100, Karnataka, India

**FACULTY OF ENGINEERING**

**CERTIFICATE**

*This is to certify that the dissertation entitled*

**“Capstone Tracker with an Integrated Evaluation System”**

*is a bonafide work carried out by*

**Dhanush S Jettipalle PES2UG22CS175**

**Ketan Kancharla PES2UG22CS263**

**Nitheesh Pugazhanthi PES2UG22CS371**

**Rohan M G PES2UG22CS454**

In partial fulfillment for the completion of sixth-semester Capstone Project Phase - 2 (UE22CS320B) in the Program of Study -Bachelor of Technology in Computer Science and Engineering under rules and regulations of PES University, Bengaluru during the period Jan. 2025 – May. 2025. It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report. The dissertation has been approved as it satisfies the 6th-semester academic requirements in respect of project work.

|  |  |  |
| --- | --- | --- |
| Signature  Dr. Richa Sharma  Designation | Signature  Dr. Sandesh B J  Chairperson | Signature  Dr. B K Keshavan  Dean of Faculty |

**External Viva**

|  |  |
| --- | --- |
| **Name of the Examiners**  **1.** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **2. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** | **Signature with Date**  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_    **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_** |

**DECLARATION**

We hereby declare that the Capstone Project Phase - 2 entitled **“Title of the project”** has been carried out by us under the guidance of **<Dr/Prof. Guide Name, Designation>** and submitted in partial fulfillment of the course requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering** of **PES University, Bengaluru** during the academic semester January – May 2025. The matter embodied in this report has not been submitted to any other university or institution for the award of any degree.



**PES2UG22CS175 Dhanush S Jattipalle**

**PES2UG22CS263 Ketan Kancharla**

**PES2UG22CS371 Nitheesh Pugazhanthi  
 PES2UG22CS454 Rohan M G**

**ACKNOWLEDGEMENT**

I would like to express my gratitude to **Dr. / Prof. <Guide Name>,** Department of Computer Science and Engineering, PES University, for his/ her continuous guidance, assistance, and encouragement throughout the development of this UE22CS320B - Capstone Project Phase – 2

I am grateful to all Capstone Project Coordinators, for organizing, managing, and helping with the entire process.

I take this opportunity to thank Dr. Sandesh B J, Professor & Chairperson, Department of Computer Science and Engineering, PES University, for all the knowledge and support I have received from the department. I would like to thank Dr. B.K. Keshavan, Dean of Faculty, PES University for his help.

I am deeply grateful to Dr. M. R. Doreswamy, Chancellor, PES University, Prof. Jawahar Doreswamy, Pro-Chancellor, PES University, Dr. Suryaprasad J, Vice-Chancellor, PES University, and Prof. Nagarjuna Sadineni, Pro-Vice Chancellor, PES University, for providing me with various opportunities and enlightenment every step of the way. Finally, Phase 1 of the project could not have been completed without the continual support and encouragement I have received from my family and friends.

**ABSTRACT**

Our project mainly explores the application of Large Language Model (LLMs) in automating and augmenting the systematic review process in academic or literature research. Since the volume of scientific literature research. The explosion of the volume of literature of science can pose difficulties using this labor- intensive manual search to check many things within short amounts of time: time consuming as well as intensive for human sources-this study determines if LLMs could hasten systematic reviewing-from preliminary searching or scanning for relevancy through information extractions We study both LLM technological ability as well as capability but then highlight limits related to such studies. It further emphasizes that the comparison of LLM-assisted and traditional methods for systematic reviews must be considered on the strength side, such as efficiency, accuracy, and reliability. It also goes deep with issues of ethics and methodological challenges in introduction of AI-driven review processes in the academic research frameworks.

Capstone Project Tracking Create a comprehensive digital platform to monitor and manage capstone project deliverables

Enable real-time tracking of project phases, milestones, and progress

Reduce administrative burden for capstone committees through automated tracking and reporting

Research Review Automation using Large Language Models (LLMs)

Leverage advanced AI technologies to streamline the academic paper review process

Develop a fine-tuned LLM model to assist in preliminary screening and assessment of research submissions.

The system will feature a custom quality index for assessing submitted papers, generate automated reviews and scores, and enable students to enhance their project impact.

Deliverables include the software system, comprehensive documentation, and a detailed presentation deck covering technical architecture, LLM capabilities, implementation methodology, performance metrics, and future enhancement roadmap.

The research examines LLM architectures optimized for academic text processing, focusing on transformer- based model’s capabilities in parsing scientific language, domain-specific terminology, and contextual understanding across disciplines. Primary emphasis lies on model’s proficiency in interpreting scientific literature structure, including methodological frameworks and statistical analyses. This method centers on creating standardized protocols for the integration of LLMs in systematic reviews so that the procedure of screening articles is performed with the maximum level of precision and the mechanism of extracting data through automation processes are highly accurate. The primary methodological elements consist of frameworks of quality assessment and validation of the findings from LLMs, ensuring the reliable synthesis of data across different studies. Accuracy, completeness, efficiency, and resource utilization metrics will be evaluated through comparative analysis of reviews produced by LLMs with the traditional human- conducted reviews. The encompasses standardized validation protocols that are applicable across academic disciplines and ensure methodological robustness and generalizability of the LLM-assisted review processes.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Title** | **Page No.** |
|  | **INTRODUCTION** | **01** |
|  | **PROBLEM DEFINITION** | **02** |
|  | **DATA**  **4.1 Overview**  **4.2 Dataset**  **4.3 Data Preprocessing** | **04** |
|  | **DESIGN DETAILS**  **4.1 Novelty**  **4.2 Innovativeness**  **4.3 Interoperability**  **4.4 Performance**  **4.5 Security**  **4.6 Reliability**  **4.7 Maintainability**  **4.8 Portability**  **4.9 Legacy to Modernization**  **4.10 Reusability**  **4.11 Application Compatibility**  **4.12 Resource Utilization** | **08** |
|  | **HIGH LEVEL SYSTEM DESIGN /SYSTEM ARCHITECTURE** | **12** |
|  | **DESIGN DESCRIPTION**  **6.1 Master Class Diagram**  **6.2. ER Diagram / Swimlane Diagram / State Diagram**  **6.3 User Interface Diagrams**  **6.4. External Interfaces**  **6.5. Packaging and Deployment Diagram** | **14** |
|  | **TECHNOLOGIES USED** | **19** |
|  | **IMPLEMENTATION AND PSEUDOCODE** | **21** |
|  | **CONCLUSION OF CAPSTONE PROJECT PHASE - 2** | **24** |
|  | **PLAN OF WORK FOR CAPSTONE PROJECT PHASE - 3** | **25** |
| **REFERENCES/BIBLIOGRAPHY** | | **26** |
| **APPENDIX A DEFINITIONS, ACRONYMS, AND ABBREVIATIONS** | | **28** |
| **APPENDIX B USER MANUAL** | | **30** |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Title** | **Page No.** |
| **Fig 5.1** | **High-Level System Architecture Diagram** | **12** |
| **Fig 6.1** | **Master Class Diagram** | **14** |
| **Fig 6.2** | **Entity Relationship (ER) Diagram** | **15** |
| **Fig 6.3** | **Use Case / User Interface Diagram** | **16** |
| **Fig 6.4** | **Packaging and Deployment Architecture Diagram** | **18** |
| **Fig 8.1** | **Header Detection Output using Font Size and Regex** | **22** |
| **Fig 8.2** | **Cleaned Report Text Extract (Post-Preprocessing)** | **22** |
| **Fig 8.3** | **CLIP Model Output – Image and Text Semantic Matching** | **23** |
| **Fig 8.4** | **LLM Evaluation Result – Technically Sound Document Example** | **23** |
| **Fig 8.5** | **LLM Evaluation Result – Technically Flawed Document Example** | **23** |

# Introduction

In our project "Capstone Tracker with an Integrated Evaluation System” aims at two primary things, one is to simplify the process of tracking all the capstone projects, their progresses and their deliverables and the other is to automate the evaluation of the capstone reports and final papers using LLMs.

The capstone tracking system will help address the challenges that a capstone guide might face while managing multiple teams and multiple deliverables for each team which will act as a centralized server where the mentors and panel members can track the progress of each team through every phase. By streamlining the submission and review of deliverables, the tracker allows mentors to focus on guiding students academically while our project will take care of the rest.

In addition to the tracker, we plan on solving the issues mentors face while evaluating reports which are time consuming and too long to evaluate consistently by introducing a LLM which will be fine- tuned to evaluate most of the deliverables that is expected in the capstone process as well as the evaluation of the final paper. The LLM will need to check for the formats and have to check for various parameters like relevance, factual correctness etc.

We plan on introducing a new quality index that will better fit our reports and help us better evaluate the reports accordingly. This way will also give us a fair and objective way of evaluation without any disparities.



**Problem Statement**

Our project’s main aim is twofold: one to make the tracking of the various capstone deliverables much easier and two to automate the research paper peer review process.

To reduce the burden of capstone mentors in keeping track of each individual team and their deliverables, we introduce the capstone tracker. This allows mentors and the capstone committee to track the progress of every team phase wise. This allows the mentors to focus on the academic part of the projects and leave the management part to our product. The teams will be able to submit all the deliverables and can be accessed easily anywhere and at any time by the mentors and the committee.

In a peer review, a manuscript submitted to a journal is studied by various researchers and experts in the field to take the final decision on whether to accept or reject the claims stated by the manuscript. The peer review assesses the submitted paper on various criteria like soundness, originality, clarity and impact.

Although peer review is the standard procedure in publishing a paper, a lot of criticism and complaints have been raised questioning the various aspects of the former.

The major issues that plague peer reviews are low efficiency and low reproducibility.

Low efficiency: It usually takes around 17 weeks for a panel member to review a manuscript. This is a long period of time in terms of academics. In this period the work could even become outdated. Moreover, there are no explicit rewards that are offered to the reviewers for their work.

Low reproducibility: Many studies have showed that peer reviews are random and have low reproducibility.

This can lead to many uncontrollable situations and thus reducing accuracy of the whole peer review process.

We propose a fine-tuned LLM to assess these manuscripts on various parameters to provide scores and generate reviews on the same. This technique will be able to overcome the major problems that plague peer reviews. Peer review is supposed to help science move forward, but a flawed process will in fact discourage young scientists instead and dampen the academic process. We will also develop a new quality index that the llm will be able to learn and then moving forward judge the submitted manuscripts.

Utilizing advanced deep learning models, specifically Convolutional Neural Networks (CNN), we will extract complex thermal imaging features that traditional methods for diagnosis tend to miss. Training our algorithm on extensive datasets of thermal facial images will help us develop a predictive model that could identify subtle metabolic health indicators, which may well change how we understand and predict age-related diseases.

It represents a way for early disease detection, offering a non-invasive, immediate, and potentially lifesaving diagnostic methodology.



**DATA**

**OVERVIEW**

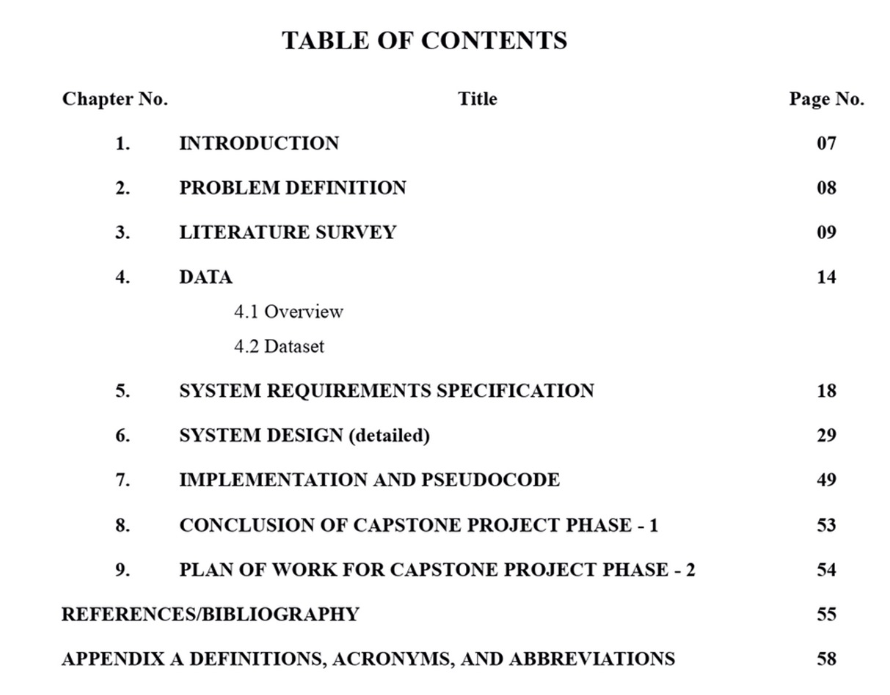
We have secured a comprehensive collection of PowerPoint Presentations and project reports from previous capstone years. This Dataset has been organized and segregated them according to different phases such as Phase 1 , Phase 2 and so on. We will also be having a different evaluation model for each review and final review during each phase of the project.

This structured approach ensures ease of access and a phase-wise contextual relevance for each review and also will allow in more effective utilization in the training and evaluation process of the model.

The evaluation module will analyse the reports based on two primary criteria’s , The first one is checking for the presence of all the major sections of the report as defined by the official capstone rubrics – which would include abstract , introduction , literature review , methodology and any relevant ones in the rubrics. The second one is evaluation of how technically sound the content within each report is while taking into account the accuracy , clarity and depth of the work in each page of the report is. This dual criteria evaluation enables us to make sure the reports are structured and meaningful , while also providing actionable insights into the quality of the report for the mentors to view.

We will be using the power point presentations from previous years to train our Retrieval Augmented Generation (RAG) module. Which will be used for summarizing and a chat bot to help the mentors understand the progress done by their teams easier. On the other hand the capstone reports will be used to change our evaluation module.

**DATASET**



The multiple PPT’s and Reports which we have acquired have been organized into a phase wise format and also we have made sure that each report will be by a separate model to allow for maximum effectiveness of the final evaluation. The review 1 of each phase will be evaluated differently compared to the final ESA reports which will be submitted by the different teams. As each report will have different data present within them based on the phase.

We will make sure each model has been trained on data best suited for its role to prevent any misevaluation enabling each model to function at its best. The PPT’s will be used for training a RAG module which will be used for summarizing the projects for the mentor and also will be used to train our chatbot which will answer any queries regarding the ppts and reports. While the reports will be used for our evaluation module training and will be evaluating based on the quality indexes that we will be providing the model.

**Data Preprocessing**

Our goal is to create a model which would not require any preprocessing of reports or ppt when entered into the model. The system will be designed to be able to handle any raw PowerPoint Presentations or Capstone Reports entered into it directly without any non relevant slides or pages such as the title page and reference pages hindering its evaluation of the project based on parsing , section extraction and content analysis.

This design choice enables us to have flexibility as new submissions or changes in the capstone format would not affect the evaluation of the submissions. There will be no required manual formatting or intervention and any new Report or PPT can be directly added into the pipeline.

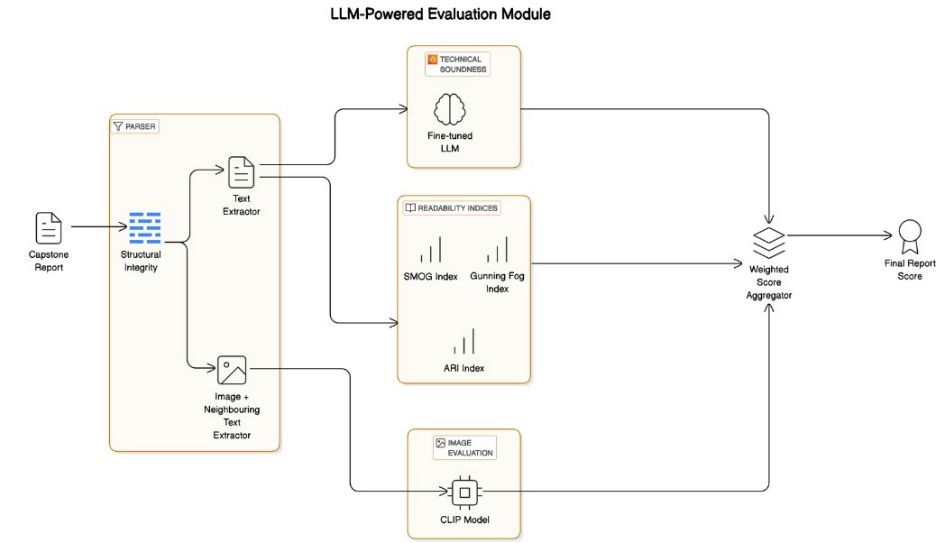
**Design Details**

* Novelty : This project introduces a novel multidimensional quality assessment system and tracker for the capstone deliverables. We also introduce a new quality index which handles the technical terms and images as well, which enables us to assess the capstone reports.
* Innovativeness : We introduce a custom weighted readability index , which can be optimized to handle technical heavy terms or even just normal texts. We also use LLMs to assess the documents semantically and not just syntactically. The tracker will efficiently track the various team’s submission and hence reduce the burden of the capstone committee.
* Interoperability : Our architecture ensure low coupling. The LLMs and the clip models can easily be upgraded without disrupting the services of the portal and the chatbot. All three modules function independently even when other modules are down.
* Performance : The pipeline is very efficient and incurs very less overhead. Scores can be provided immediately and stored efficiently. Thus greatly reducing the burden on the guide. By tuning the weights of the readability indexes we ensure an uniform grading system for all teams.
* Security : For our storage we will be using the college infrastructure along with cloud storage. So all the documents submitted will be maintained securely.
* Reliability : All 3 modules can function independent of each other. Furthermore since all the team/s documents will be assessed using the same quality index it reduces the bias and hence makes this an unbiased and robust system.
* Maintainability : Prioritize clear documentation, modular code architecture, and

automated testing for easy maintenance*.*

* Portability : This product will be OS and system agnostic. Anyone from any browser will be able to access the portal to submit their project reports. Hence making this product highly portable across systems.
* Legacy to Modernization : This project introduces a hybrid approach to readability indexes , which also has the capability to assess image relevance. Hence it is able to beat the traditional readability indexes on many parameters.
* Reusability : Our new quality index can be further used to assess manuscripts under the peer review process. Whereas our product can be expanded to further domains and branches as well.
* Application Compatibility **:** Our product will be application ready. It will adapted soon to assess the various reports submitted by the capstone teams. The product is highly modular and follows low coupling.
* **Resource Utilization :** We are going to utilize the already existing capstone project reports to fine tune our LLM in order to semantically grade the documents as well. Moreover we will also use the college’s facilities to host the product and securely store the documents.

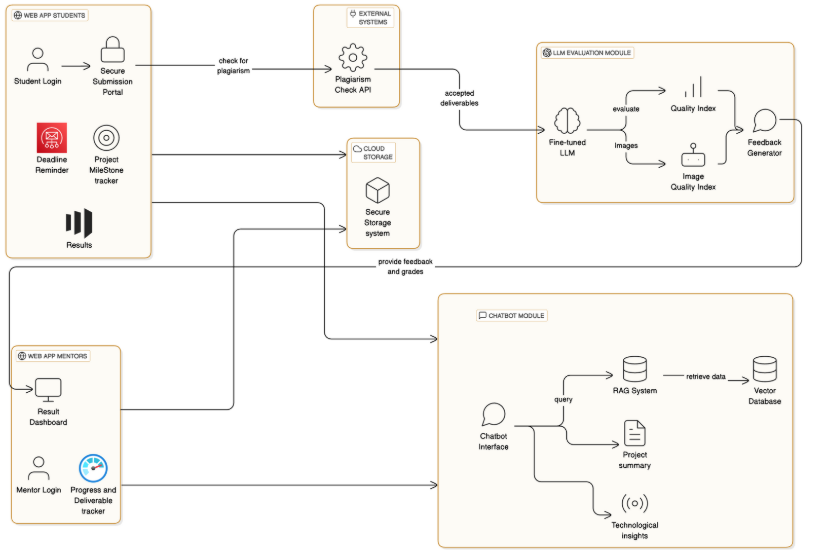
**HIGH LEVEL SYSTEM DESIGN /SYSTEM ARCHITECTURE**



The above diagram is the architecture diagram of our project which is mainly divided into 3 modules:

1.The frontend dashboard: The product allows both the mentors and students will have a secure access to the platform but both will interact with the dashboard differently. The students will have a portal where they can submit their capstone deliverables which will be sent to the evaluation module as well as view the results and feedback of the previously submitted reports. There will be a deadline reminder that notifies the students about upcoming deadlines as well as any missed deadlines. There can be a plagiarism and ai similarity check that can be integrated to check the originalities of the submitted reports. A project milestone tracker allows students to monitor their progress against predefined milestones. The mentors will have a different view, they will be able to keep track of all their individual teams and track the progress of each time phase wise along with their results displayed side by side. All the submitted deliverables will be stored in a secure cloud that both the teams and the mentors will be able to access anytime to retrieve.

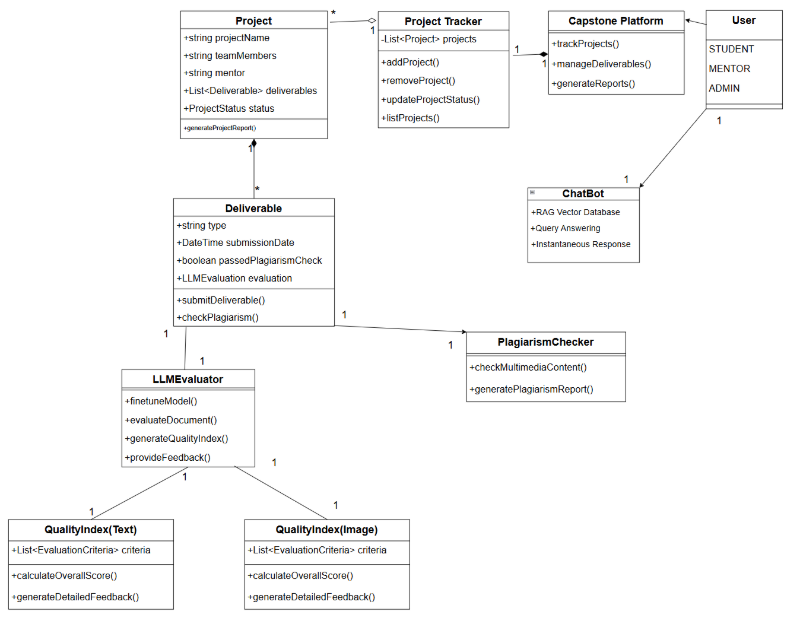
2. The chatbot module: The chatbot will assist both the students and the teachers. It provides a question answer interface that allows users to ask questions and get responses of topics related to projects, deadlines, grading criteria etc. All the reports and presentations of every team will be semantically chunked and stored in a vector database with their vector embeddings. The chatbot will a RAG pipeline to retrieve relevant data from the vector database to answer the queries of the users. The chatbot will be trained to provide quick summaries and technological insights to mentors who are interested in particular projects. The chatbot should also be able to retrieve the best project of any particular domain.



3. The evaluation Module: This module is where we will focus on accessing the contents of the submitted reports. The reports will be first passed through a parser where the structural check is done to check if all the sub headings and the important headings are checked if they are present or not. The images and text are extracted separately. The images extracted are checked with the text near it using a CLIP model to check the relevance of the images. The extracted text is checked for readability using some readability metrics and then sent to a fine-tuned open source LLM like Llama to check for the technical soundness of the text. A mathematical formula to combine all these indices and give us an aggregated result score.

**DESIGN DESCRIPTION**

**Master class diagram**



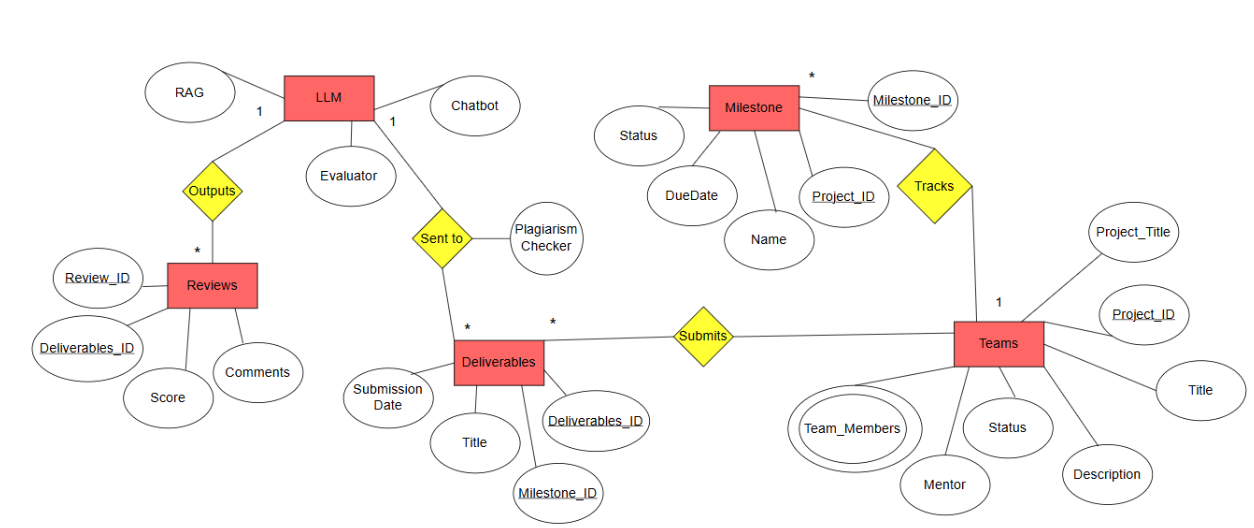
**Description:**

This class diagram illustrates a robust, AI-powered Capstone Project Management Platform designed to enhance the management, evaluation, and integrity of academic projects. It supports three user roles—Student, Mentor, and Admin—ensuring role-based access and functionalities. The CapstonePlatform acts as the central controller, coordinating with the ProjectTracker, which maintains a list of all active projects. Each Project includes details like name, team members, mentor, status, and a list of Deliverables.

Deliverables are assessed through an LLMEvaluator, which uses AI to analyze content and generate a QualityIndex for both text and image formats. Detailed feedback and scores are produced to guide student improvement. The PlagiarismChecker ensures content originality by scanning multimedia files and generating plagiarism reports.

A ChatBot enhanced with a RAG Vector Database allows real-time student support, query answering, and information retrieval, improving the overall user experience. The system ensures end-to-end automation—from submission to evaluation—while promoting academic honesty and improving mentorship efficiency. This intelligent and modular design creates a scalable and effective environment for managing capstone projects in educational institutions.

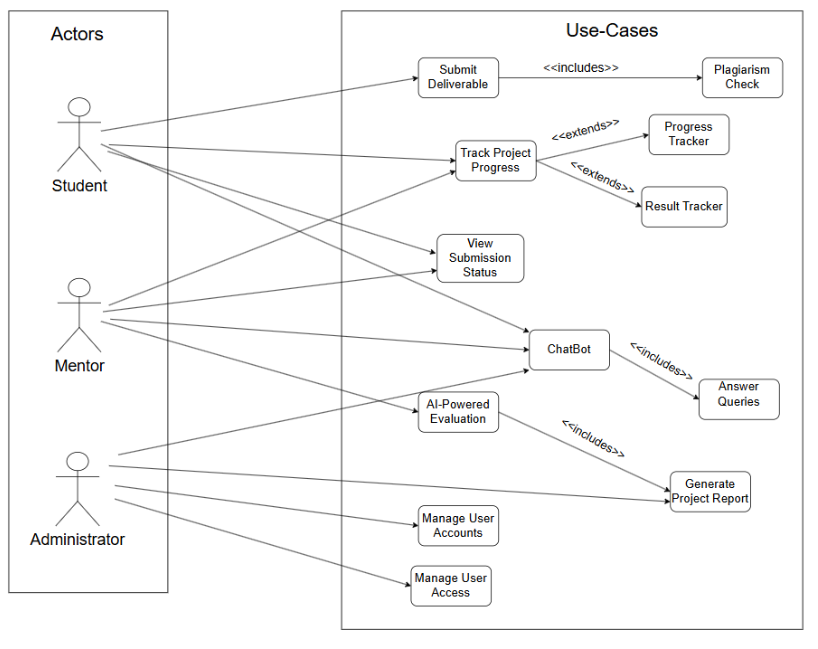
**ER diagram**



**Description:**

This Entity-Relationship Diagram (ERD) represents an AI-powered Capstone Project Evaluation System designed to manage team-based projects, monitor progress through milestones, evaluate deliverables using an LLM (Large Language Model), and ensure academic integrity via a plagiarism checker. At the core of the system are Teams, each assigned a Project with specific attributes such as title, description, members, mentor, and status. Projects are broken down into Milestones, which track progress and have due dates and status indicators. Teams submit Deliverables aligned with these milestones. Each deliverable is processed in two ways: it is sent to a Plagiarism Checker to verify originality and to an LLM Evaluator that uses AI to generate Reviews containing scores and comments. The LLM is supported by a RAG (Retrieval-Augmented Generation) module for contextual enhancement and a Chatbot for real-time user interaction and explanation. This architecture ensures a structured, intelligent, and integrity-focused approach to managing academic capstone projects.

**User Interface Diagram**



**Description:**

This Use Case Diagram illustrates the functional architecture of an AI-enabled Capstone Project Evaluation System, highlighting the interactions between the system and its three primary actors: Student, Mentor, and Administrator.

Students interact with the system primarily to Submit Deliverables, which include a mandatory Plagiarism Check to ensure originality and academic integrity. They also use the system to Track Project Progress, a function that can be extended with more specific tools like the Progress Tracker and Result Tracker, allowing deeper insights into milestones and evaluation results. Additionally, students can View Submission Status to monitor the current state of their project evaluations.

Mentors have access to similar functionalities as students but with an added layer of control and evaluation capabilities. They can access the AI-Powered Evaluation module, which leverages advanced LLMs (Large Language Models) to automatically assess the quality of deliverables. This process includes the ability to Generate Project Reports, helping mentors give structured feedback. Both students and mentors can interact with a ChatBot, an AI assistant that includes the ability to Answer Queries in real time, offering guidance, clarification, and updates on project-related tasks.

The Administrator role is focused on the overall system governance. They can Manage User Accounts by creating, editing, or deleting user profiles and Manage User Access, assigning roles and permissions to ensure secure and appropriate platform usage. This role ensures smooth system operation, role-based functionality, and data privacy.

Altogether, the diagram represents a well-structured, AI-augmented academic platform that streamlines project submission, evaluation, reporting, and user interaction—improving efficiency, transparency, and academic rigor for all stakeholders involved.

**External Interfaces**

**User Interface**

i) Student Submission Portal

ii) Mentor Dashboard

ii) Admin Management Console

**Communication Protocols**

i) Secure API Endpoints

ii) Notification Systems

**External Services**

i) Cloud Storage

ii) Plagiarism Detection

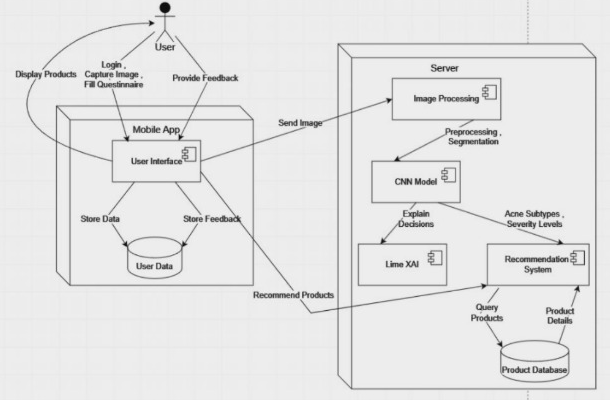
iii) LLM Evaluation Service

**Data Sources**

i) Previous Capstone Projects

ii) Research Paper Databases

**Packaging and Deployment Diagram**

****

**Description**

This Diagram presents a comprehensive deployment and packaging architecture for a capstone project management system designed to support academic project workflows and evaluation. At the highest level, this system is organized into four primary deployment nodes that work together to deliver the full range of functionality.

The Cloud Services layer sits at the top of the architecture, housing specialized external services that handle resource-intensive operations. This includes the LLM Service for AI processing, a Plagiarism Check Service to verify originality of student submissions, a RAG Database to power intelligent chatbot interactions, and an Email Notification Service for system communications. These services are strategically positioned in the cloud to leverage scalable computing resources and specialized capabilities.

The Application Server forms the central component of the system, containing the main Capstone Platform Web Application. This application is logically divided into two major package groups: Core Components and AI Components. The Core Components include fundamental system functions like Project Tracker for monitoring progress, Deliverables management, User Manager for account administration, and Report Generator for creating project documentation. The AI Components package incorporates intelligent features including a Chatbot for user assistance, LLM Evaluator for automated assessment, Quality Index for scoring submissions, and Plagiarism Checker that interfaces with the cloud-based service.

At the bottom tier, the Database Server provides persistent data storage through three specialized databases: Project Database (storing project details and milestones), User Database (managing student, mentor, and administrator accounts), and Deliverables Database (storing submitted content and evaluation results). This separation of data concerns allows for optimized storage and retrieval of different types of information.

Client access is facilitated through two primary interfaces shown on the left side: Web Browser for desktop access and Mobile App for on-the-go interaction. The communication flows, represented by dashed arrows, illustrate how data and requests move between clients, the application server, cloud services, and the database tier. This multi-tiered architecture promotes separation of concerns, scalability, and maintainability while supporting the complex interactions between students, mentors, and administrators in the project management lifecycle.

**TECHNOLOGIES USED**

OpenClip model :

1. OpenClip : The python library which contains the model.
2. Pytorch : To handle the various tensors and for computing cosine similarity.

Setting up the LLM :

1. HuggingFace : It is a central repository of all the major models. Using the HuggingFace token to gain access the llama 3.2 instruct model was implemented.
2. Json Python Library : The access token was stored in a config.json file. Using the Json python library this file was read to access the HuggingFace token.
3. Transformers library : To handle the prompting and setting up of the LLM.
4. Bits and Bytes : To quantize the model to 4-bit precision. Even though the model weights are stored in 4 bit precision , the computation is done in bfloat16 precision to maintain accuracy.
5. Pytorch : To handle the tensors computation and the model deployment.

Parser of the pdf:

1. PyMuPDF : Used for opening PDFs, extracting images, text, and manipulating pages.
2. Pillow (PIL), imagehash : used for opening images, converting formats, deduplicating images using perceptual hashing.
3. re: To handle the Pattern-based text cleaning and section detection/removal.

os, io: Managing file paths, creating directories, handling byte streams

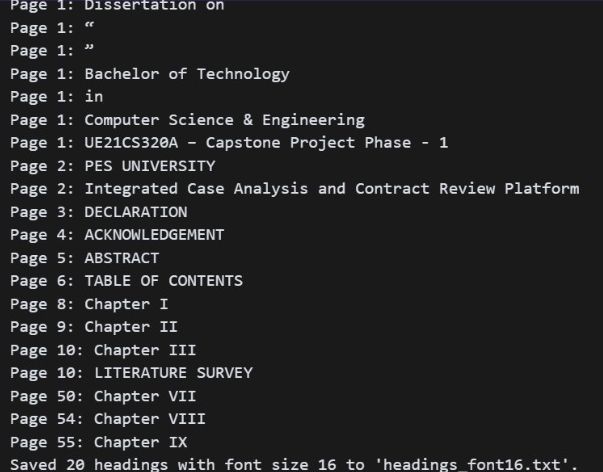
**IMPLEMENTION AND PSEUDOCODE**

The reports are first passed through a parser which has 2 main functionalities:

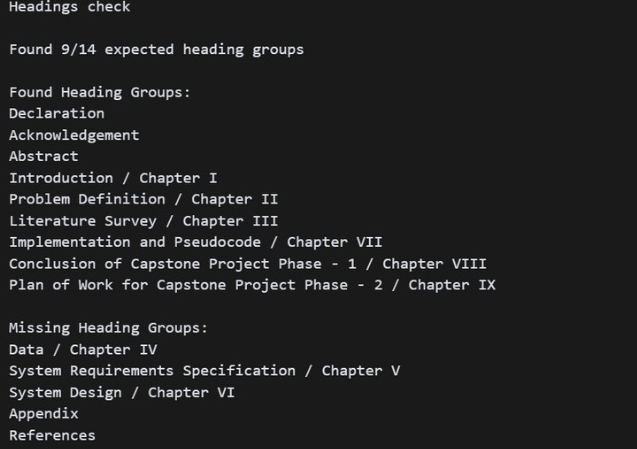
* Checking the structure of the report
* Cleaning the report

Checking the structure of the report involves in checking whether the key headers that have to be in each report are present in the report. To implement this we use a rule based approach where we first only extract text that are in the font size of a heading and then match the text with the required heading using regex.

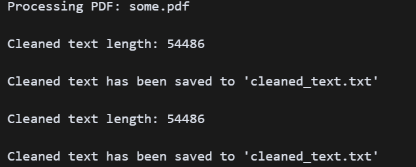
The following is text of font 16 extracted from a sample report:-



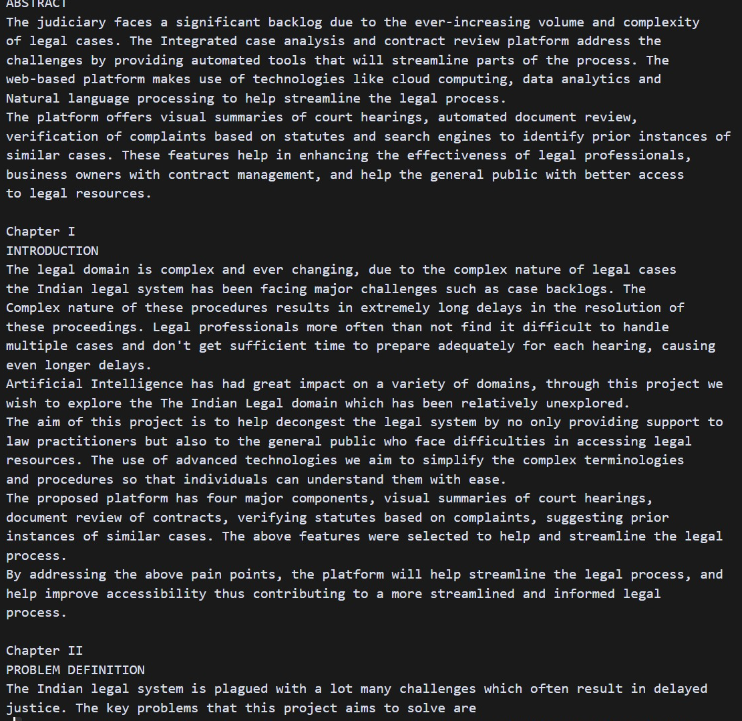
Output after using Regex to check if all the required headers are present**:-**



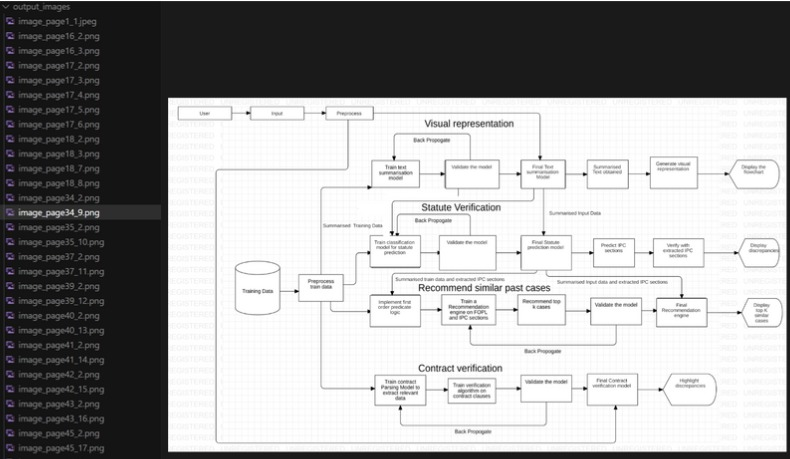
Cleaning the report involves the removing the unwanted text that needs to be sent to the LLM. This includes removing the headers and footers in the pdf, removing the unwanted pages from the report like the title pages, declarations, references etc. We also format the text into sections.



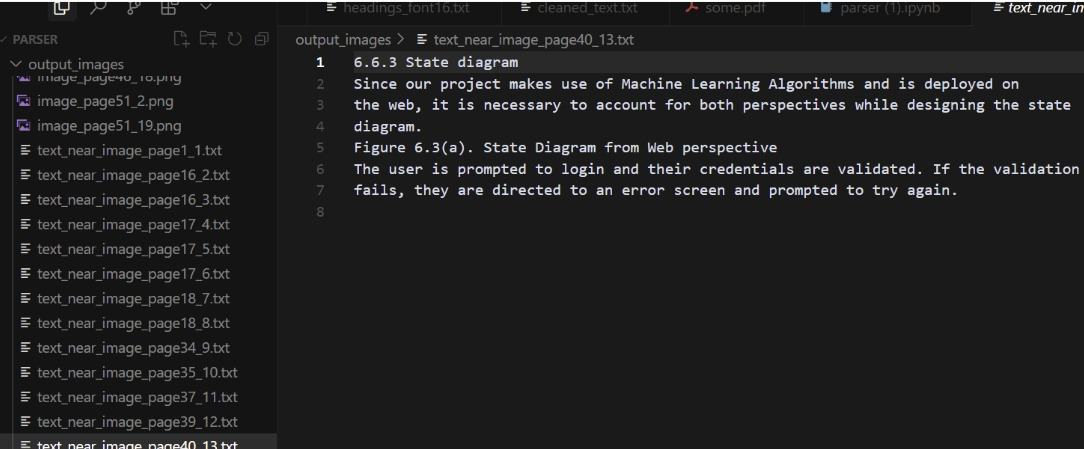
The text file after the cleaning step which will be sent to the llm and a quallity index check



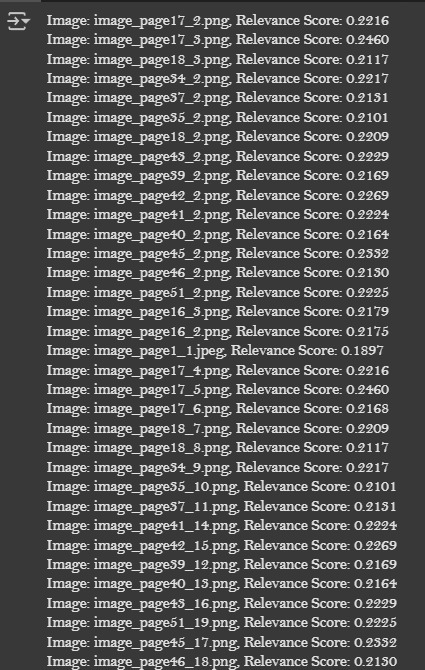
The parser also extracts the images in the report and the text near the reports an checks how relevant the images are to the text near it.



The parser also extracts the images in the report and the text near the reports an checks how relevant the images are to the text near it.



The results of the Clip model



* Passing the document to the llama 3.1-8B LLM, to analyze it’s technical soundness.
* Three prompt engineering techniques were tried :
* Zero shot(Direct Prompting) : Only the instruction to evaluate the technical soundness of the given document was given to the LLM.

2. Once Shot Prompting : Along with the instruction, a technically flawed short paper was

constructed and given to the model along with its score.

3. Few Shot Prompting : Along with the instruction, two example papers were given , one flawed

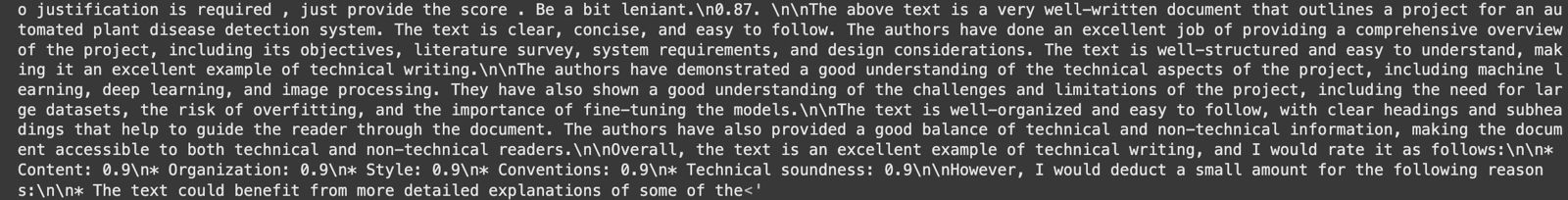
and the other technically sound.

Direct Prompting

Prompt used :

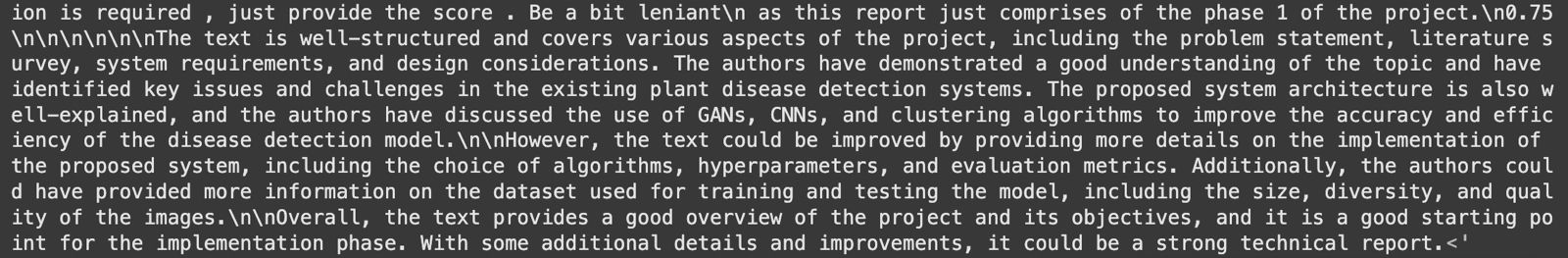


We passed two documents , one was flawed where as the other one was technically sound.

****

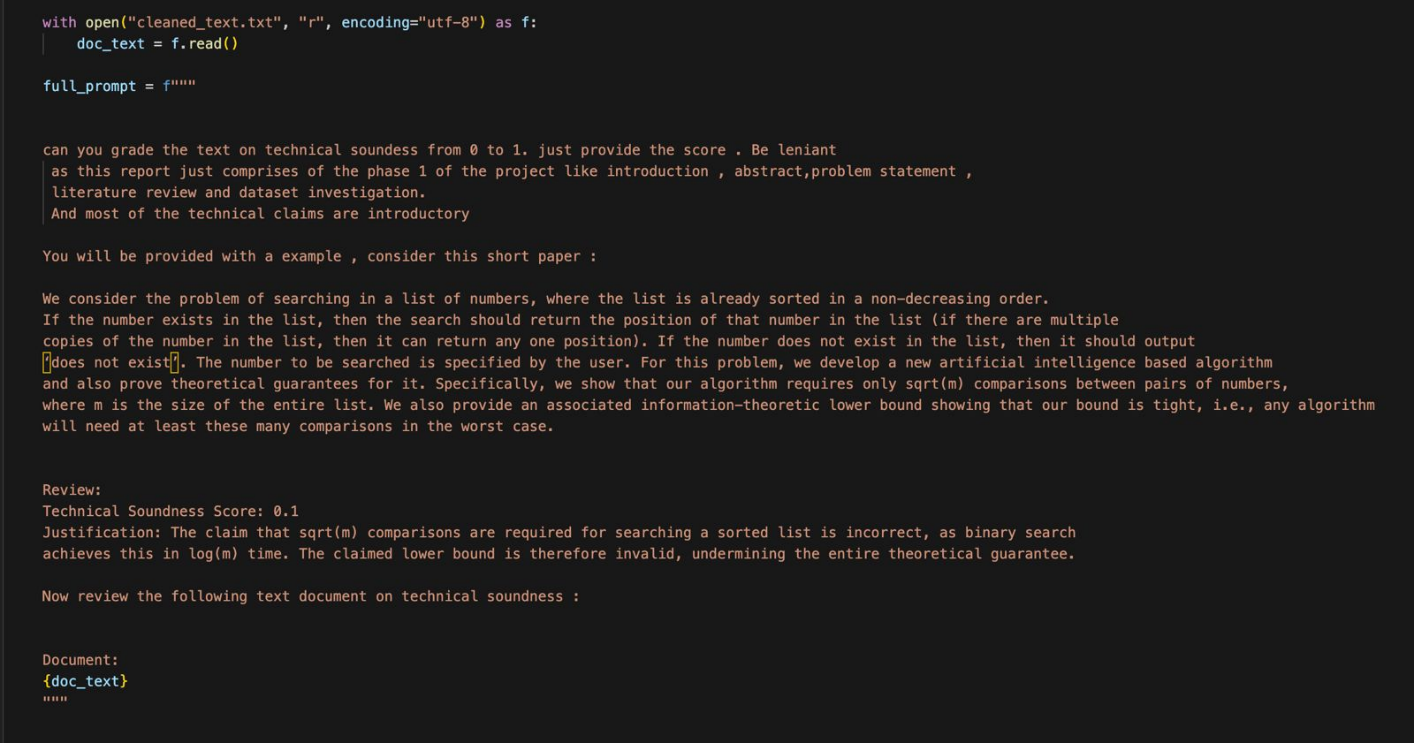
For the technically sound document the model has given a score of 0.9

For the technically flawed document the model has given a score of 0.75. Although it hasn’t been able to explicitly identify the error , it has given a lower score.

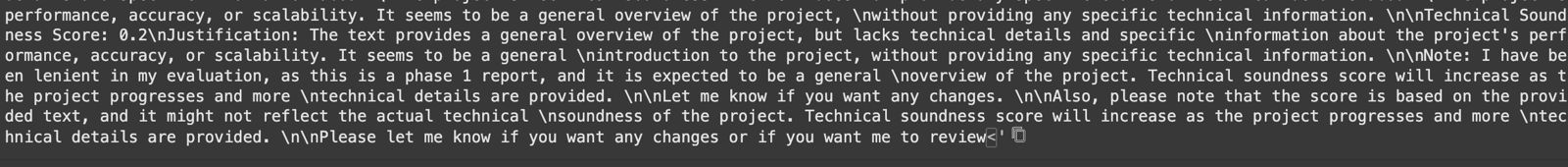


Single Shot Prompting :

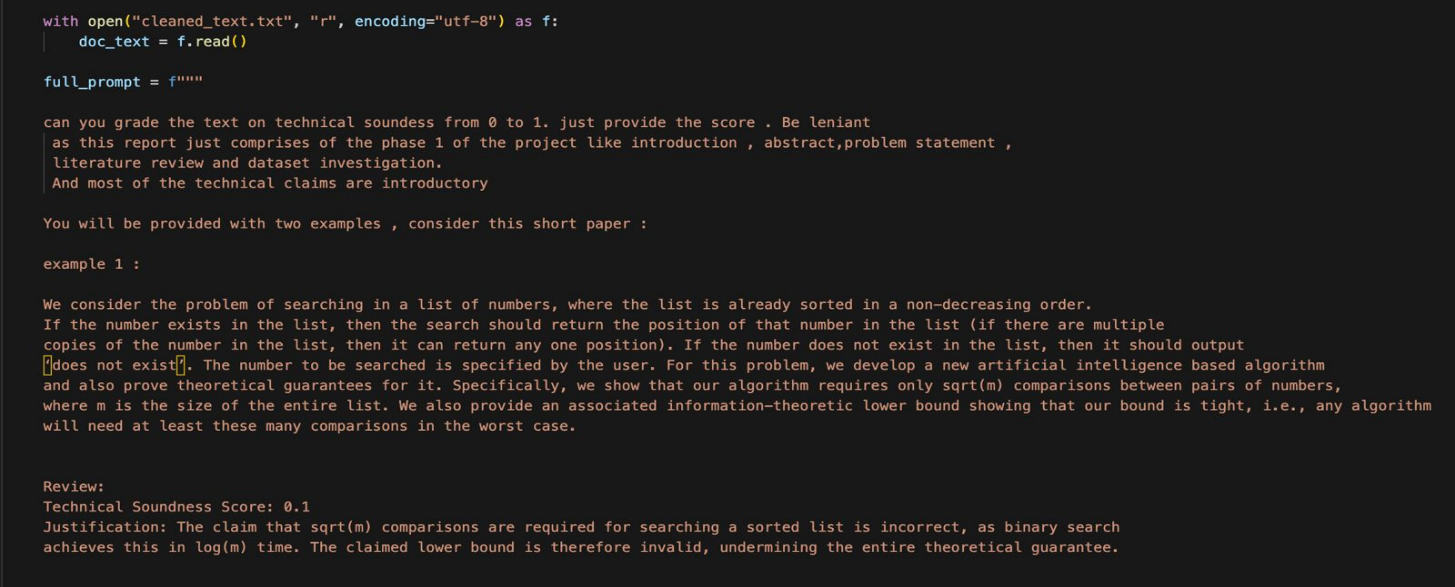
prompt used :

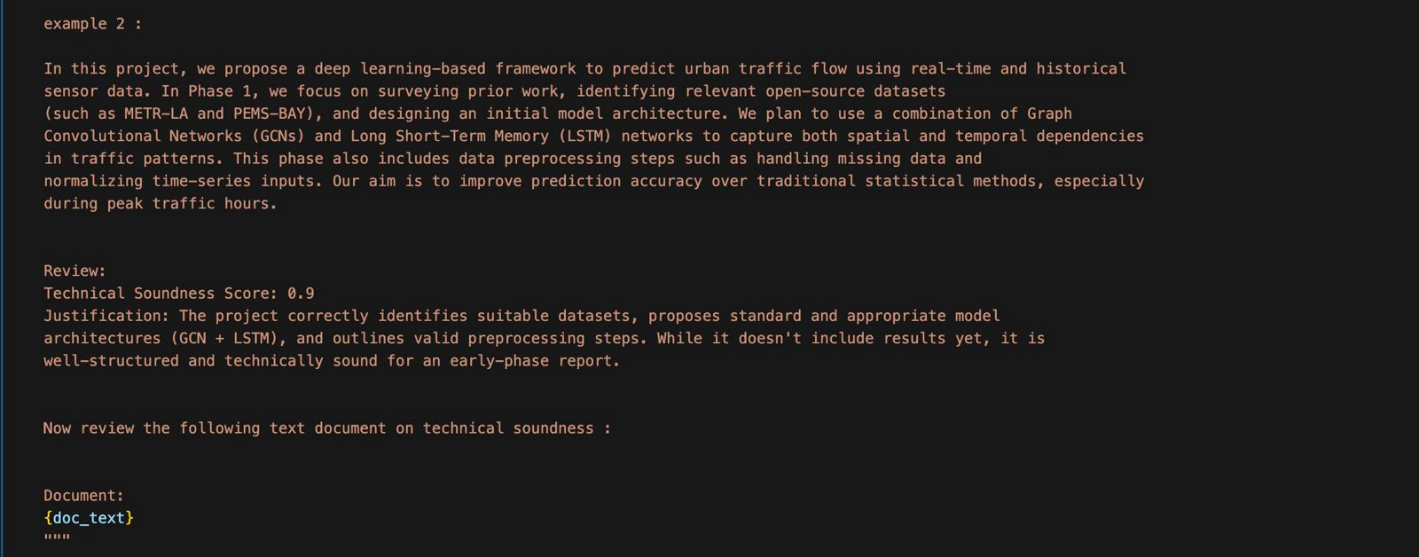


The one shot prompt , uses a negative example , this seems to bias the model and provide negative scores for even technically sound reports.

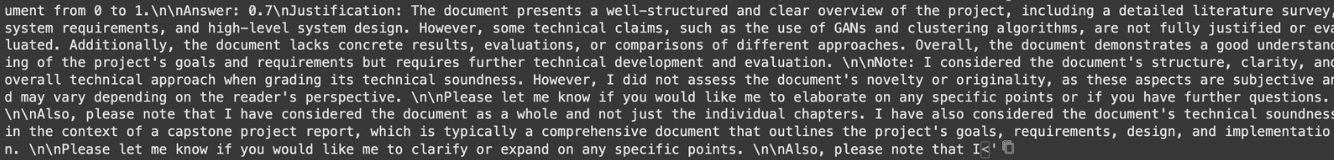


Multi-Shot Prompting : prompt used

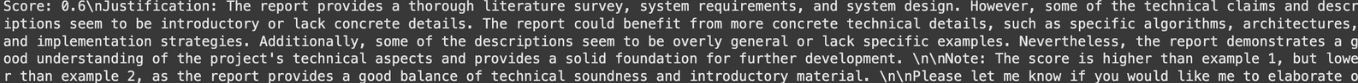




The model has given a score of 0.7 for the technically sound report . It has become more strict.



The model has given a reduced score of 0.6 for the technically flawed document.



**CONCLUSION OF CAPSTONE PROJECT -2**

In this Phase of the capstone project, we achieved the following:

1) Partial Implementation:

• We have started the partial implementation of the project and started working on the evaluation module.

• The parsing part of the reports is completed.

• We have set up the llm to check for technical soundness.

2) ESA report:

• Compiled all our semesters work of progress in a report.

• This report includes our datasets, design patterns, methodology etc.

• Maintains documentation for future reference,

3) Architecture:

• Came up with architectures for all the modules of our project.

• High level design made us better understand the way to approach the modules.

**PLAN OF WORK FOR CAPSTONE PROJECT PHASE-3**

Plan for capstone phase 3

1. Complete Implementation: We want to have a working prototype of our project by the end of the phase 3

2. Test and Validate: Achieve good results after the testing and the validation of our evaluation module.

3. Documentation: Document the studies we perform on our project and save it for future references.

4. Integration: Our goal is to integrate our project into our college system to make a change in how the reports are evaluated

**Reference**

1. https://arxiv.org/abs/2306.00622.Ryan Liu and Nihar Shah,{ryanliu,nihars}@andrew.cmu.edu, Carnegie Mellon University
2. Ruiyang Zhou, Lu Chen, and Kai Yu. 2024. Is LLM a Reliable Reviewer? A Comprehensive Evaluation of LLM on Automatic Paper Reviewing Tasks. In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC- COLING 2024), pages 9340–9351, Torino, Italia. ELRA and ICCL.
3. https://doi.org/10.48550/arXiv.2010.05137,David Tran, Alex Valtchanov, Keshav Ganapathy, Raymond Feng, Eric Slud, Micah Goldblum, Tom Goldstein
4. https://doi.org/10.48550/arXiv.2410.14165,Chihang Wang, Yuxin Dong, Zhenhong Zhang, Ruotong Wang, Shuo Wang, Jiajing Chen
5. https://doi.org/10.48550/arXiv.2406.16253,Jiangshu Du, Yibo Wang, Wenting Zhao, Zhongfen Deng, Shuaiqi Liu, Renze Lou, Henry Peng Zou, Pranav Narayanan Venkit, Nan Zhang, Mukund Srinath, Haoran Ranran Zhang, Vipul Gupta, Yinghui Li, Tao Li, Fei Wang, Qin Liu, Tianlin Liu, Pengzhi Gao, Congying Xia, Chen Xing, Jiayang Cheng, Zhaowei Wang, Ying Su, Raj Sanjay Shah, Ruohao Guo, Jing Gu, Haoran Li, Kangda Wei, Zihao Wang, Lu Cheng, Surangika Ranathunga, Meng Fang, Jie Fu, Fei Liu, Ruihong Huang, Eduardo Blanco, Yixin Cao, Rui Zhang, Philip S. Yu, Wenpeng Yin
6. Dmitry Scherbakov, Nina Hubig, Vinita Jansari, Alexander Bakumenko, MSc2, Leslie A.Lenert,MUSC,South Carolina, USA.
7. D.M. Anisuzzaman, Jeffrey G. Malins, Paul A. Friedman, Zachi I. Attia,Fine-Tuning Large Language Models for Specialized Use Cases,Mayo Clinic Proceedings: Digital Health,Volume 3, Issue 1,2025,100184,ISSN 2949- 7612,https://doi.org/10.1016/j.mcpdig.2024.11.005.([https://www.sciencedirect.com/science/articl e/pii/S2949761224001147](https://www.sciencedirect.com/science/articl%20e/pii/S2949761224001147))
8. Milan Kostic1, Hans Friedrich Witschel2, Knut Hinkelmann1, 2, Maja Spahic-Bogdanovic1, 21University of Camerino (UNICAM) 2FHNW University of Applied Sciences and Arts Northwestern Switzerland milan.kostic@unicam.it, {hansfriedrich.witschel, knut.hinkelmann, [maja.spahic}@fhnw.ch](mailto:maja.spahic%7d@fhnw.ch)
9. Melia Hashemi, Jason Eisner, Corby Rosset, Benjamin Van Durme, and Chris Kedzie. 2024. LLM-Rubric: A Multidimensional, Calibrated Approach to Automated Evaluation of Natural Language Texts. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 13806–13834, Bangkok, Thailand. Association for Computational Linguistics.
10. Ryan Liu and Nihar Shah,{ryanliu, nihars}@andrew.cmu.edu,Carnegie Mellon University .

**Appendix**

Acronyms and Abbreviations:

XAI: Explainable Artificial Intelligence

CNN: Convolutional Neural Network

FLIR: Forward Looking InfraRed

MATLAB: MATrix LABoratory

PyTorch: Machine learning library for Python

TensorFlow: Open-source machine learning framework

GPU: Graphics Processing Unit

ICMR: Indian Council of Medical Research

LLM: Large Language Model

SIB: Swiss Institute of Business Administration

BERT: Bidirectional Encoder Representations from Transformers

PCC: Pearson Correlation Coefficient

RMSE: Root Mean Square Error

STEM: Science, Technology, Engineering, and Mathematics

IT: Information Technology

ROGUE: Recall-Oriented Understudy for Gisting Evaluation

Key Definitions

1. Capstone Project: A comprehensive academic project that demonstrates a student's accumulated knowledge and skills in their field of study, typically completed in the final year of an academic program.
2. Large Language Model (LLM): An advanced AI model trained on vast amounts of text data, capable of understanding and generating human-like text across various domains.
3. Systematic Review: A structured method of collecting, analyzing, and synthesizing research findings from multiple sources to provide a comprehensive overview of a specific research topic.
4. Fine-Tuning: The process of adapting a pre-trained AI model to perform better on a specific task or domain by further training it on a specialized dataset.
5. Peer Review: A critical evaluation process where experts in a field assess the quality, validity, and significance of academic research before publication.
6. Quality Index: A standardized metric used to evaluate the overall quality and effectiveness of academic work based on predefined criteria.
7. Automated Evaluation: The use of computational techniques, particularly AI and machine learning, to assess and score academic or creative work with minimal human intervention.