

INDUSTRIAL ATTACHMENT REPORT CARRIED OUT AT RP MUSANZE COLLEGE FROM 9th SEPTEMBER 2024 UP TO 5th NOVEMBER 2024

Company name: RP MUSANZE COLLEGE

Company supervisor's name: NIYITEGEKA Jean Joas

Phone number: +250 788433314

Name of student: NIYOGUSHIMWA Natanael

Registration number: 22RP08913

Telephone number: 0781796824

Email: nathanaelniyogushimwa@gmail.com

Department: Information Communication Technology

Option: Information Technology

Academic year: 2023-2024

DECLARATION

I hereby declare that I successfully completed an industrial attachment at **RP Musanze College**

where I engaged in various activities aligned with the goals and objectives of the organization.

IA was undertaken as part of my educational and professional development to gain practical

experience in the field.

During this period, I adhered to the rules, regulations, and policies of RP Musanze College,

contributing to projects and assignments assigned to me. I gained valuable skills, knowledge, and

insights that will be instrumental in my future professional pursuits.

I, therefore, declare that this work is my own contribution to the partial fulfilment of the award

of the degree of diploma in ICT department, in Information Technology Program at Integrated

Polytechnic Regional College MUSANZE.

Date:/2024

Signature:

Names: NIYOGUSHIMWA Natanael

APPROVAL

This is to certify that the trainee called NIYOGUSHIMWA Natanael has completed his industrial attachment program from 9th September up to 5th November 2024 in RP Musanze College,

under our supervision in Information Communication Technology (ICT).

Company Supervisor:
NIYITEGEKA Jean Joas
Date :/2024
Signature

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to **RP Musanze College** for providing me with the opportunity to complete my industrial attachment. The experience has been invaluable for both my professional and technical growth, as it allowed me to gain hands-on skills in several areas of ICT and technical support.

My deepest appreciation goes to my supervisors and the entire technical team at RP Musanze College for their guidance, support, and encouragement throughout the attachment.

DEDICATION

I humbly dedicate this industrial attachment to my Almighty God for all his guidance, protection, help and his Love along the time of industrial attachment and support enabled me to gain practical experience and confidence in the field.

This achievement would not have been possible without their guidance and the hands-on experience I gained during my time at RP Musanze College.

I also dedicate this work to **NIYITEGEKA Jean Joas** for his ideal support

TABLE OF CONTENTS

DECLARATION	i
APPROVAL	ii
ACKNOWLEDGEMENT	iii
DEDICATION	iv
TABLE OF CONTENTS	v
LIST OF FIGURES	vi
LIST OF ABBREVIATIONS AND ACRONYMS	vii
DEFINITION OF TERMS	viii
EXECUTIVE SUMMARY	ix
CHAPTER 1: GENERAL INTRODUCTION	
Organizational Profile:	1
CHAPTER 2: PERFORMED ACTIVITIES IN INDUSTRIAL ATTACHMENT	2
PROGRAM	
GENERAL ACTIVITIES	
SPECIFIC ACTIVITIES	
Analysis of Learned Knowledge and Applied Skills:	4
A Profile of Skills and Competencies Gained/Acquired:	5
CHAPTER 3: INDUSTRIAL ATTACHMENT LIFE-EXPERIENCE	7
CHAPTER 4: GENERAL CONCLUSION AND RECOMMENDATIONS	8
Summary:	8
Conclusion and Recommendations:	9
Conclusion	9
Recommendations	10
REFERENCES	10
APPENDICES	11



LIST OF FIGURES

Figure 1 – Organization structure	2
Figure 2	12
Figure 3	12
Figure 4	13
Figure 5	14
Figure 6	14
Figure 7	15
Figure 8	16
Figure 9	17
Figure 10	17
Figure 11	18
Figure 12	19
Figure 13	19
Figure 14	20
Home Page1 Figure 15	20
Home page2 Figure 16	21
Project page Figure 17	21
Admin page1 Figure 18	22
Admin Page2 Figure 19	22
Figure 20 – Admin Saving project page	23
Figure 21 – Recorded Project with no plagiarism detected.	23
Figure 22 – Project overview page	24
Figure 23 – Plagiarism checker Home page	24
Figure 24 – output 1	25
Figure 25 – output 2	26
Figure 26 – output 3	26







LIST OF ABBREVIATIONS AND ACRONYMS

- ❖ AI: Artificial Intelligence
- ❖ BERT: Bidirectional Encoder Representations from Transformers
- * RP: Rwanda Polytechnic
- ❖ NLP: Natural Language Processing
- ❖ DMS: Document Management System
- ORM: Object-Relational Mapping
- ❖ SQL: Structured Query Language
- UI: User Interface
- ❖ API: Application Programming Interface

DEFINITION OF TERMS

✓ Plagiarism:

The act of using someone else's work, ideas, or text without proper acknowledgment, presenting it as one's own. In academic research, plagiarism is considered a serious violation of academic integrity.

✓ Artificial Intelligence (AI):

The simulation of human intelligence processes by machines, particularly computer systems. AI involves the development of algorithms and models that enable computers to perform tasks such as learning, reasoning, and problem-solving.

✓ BERT (Bidirectional Encoder Representations from Transformers):

✓ A pre-trained natural language processing (NLP) model designed to understand the context of words in a sentence. BERT is used for various text analysis tasks, including sentiment analysis, text classification, and similarity detection.

✓ Natural Language Processing (NLP):

A field of AI that focuses on the interaction between computers and human language. NLP techniques enable computers to process, analyze, and understand large amounts of natural language data.

✓ Semantic Analysis:

The process of understanding the meaning and context of text by analyzing word relationships and context. In the context of plagiarism detection, semantic analysis helps identify paraphrased content that may not have exact word matches but has similar meanings.

✓ Embedding:

A numerical representation of text data in a multi-dimensional space, where similar words or sentences have similar embeddings. Embeddings are used in the system to calculate similarity scores between documents.

✓ Django Framework:

A high-level Python web framework that enables the rapid development of secure and maintainable web applications. Django simplifies tasks such as database management, URL routing, and user authentication.

✓ Document Management System (DMS):

A software system used to store, manage, and track electronic documents. In this project, the DMS is integrated with the anti-plagiarism system to securely store and manage unique academic research projects.

✓ Similarity Score:

A numerical value that indicates how similar two documents are based on their content. The score is calculated using text embeddings and helps determine whether an uploaded document contains plagiarized content.

✓ Preprocessing:

The process of cleaning and preparing text data for analysis. Preprocessing may include tasks like tokenization, stopword removal, and stemming, which help improve the accuracy of text analysis.

EXECUTIVE SUMMARY

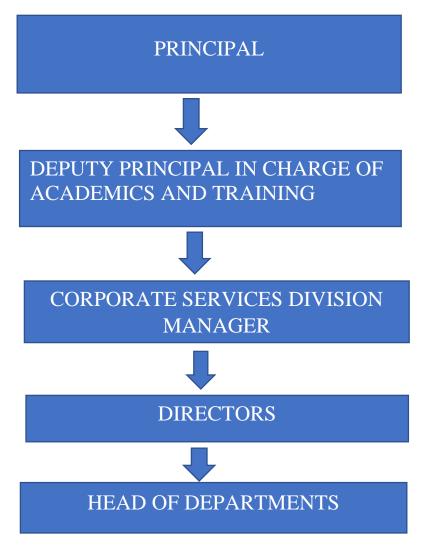
The industrial attachment at RP Musanze College focused on the development of an Anti-

Plagiarism System for academic research projects. The system employs Artificial Intelligence (AI) and the BERT model to detect plagiarism in academic projects uploaded by final-year students. The project involved designing and implementing the AI system in Python using Django for web development, with a focus on ensuring data integrity, security, and scalability. This report outlines the work done, the skills gained, and the impact of the system on academic integrity.

CHAPTER 1: GENERAL INTRODUCTION Organizational Profile: ➤ Geographical Location and Historical Background: RP Musanze College is located in Musanze, Northern Rwanda. It is part of the Rwanda Polytechnic (RP) network, which aims to provide high-quality technical and vocational training. > Organizational Structure: RP Musanze College is structured with various departments including Information Technology, Civic Engineering, Electrical and Electronic Engineering, Hospitality

Management and Agriculture Engineering. Each department has specific roles and responsibilities to ensure the effective delivery of academic programs.

Figure 1 – Organization structure



Details of the Placement Department:

The IT department of RP Musanze College is responsible for training students in various IT fields, including software development, networking, and database management. During my attachment, I was placed in the department where I worked on AI and web development.

CHAPTER 2: PERFORMED ACTIVITIES IN INDUSTRIAL ATTACHMENT PROGRAM

GENERAL ACTIVITIES

Collaborating with the IT Department:

Worked closely with supervisors and team members in the IT department, participating in regular meetings and discussions to understand project requirements and provide updates on progress.

Research and Familiarization with Plagiarism Detection Techniques:

Conducted research on state-of-the-art plagiarism detection methods and familiarized myself with machine learning models like BERT, focusing on their application in text similarity analysis.

Setting Up the Development Environment:

Assisted in setting up the necessary development environment, including configuring Python, Django, and relevant machine learning libraries, to support the anti-plagiarism system.

Learning and Applying Python and Django for Web Development:

Gained hands-on experience in Python programming and web application development using the Django framework, focusing on both backend logic and frontend integration.

Problem Solving and Debugging:

Engaged in problem-solving and debugging activities to address issues encountered during the implementation phase, improving system performance and stability.

SPECIFIC ACTIVITIES

Designing the Database Schema for Storing Student Projects:

Created a detailed database schema using Django's ORM, designed to store project metadata (title, author, date) and securely manage uploaded documents for plagiarism checks.

4 Implementing the AI-Based Plagiarism Detection Algorithm:

Developed a plagiarism detection algorithm using the BERT model for semantic analysis, enabling the system to identify both direct copying and paraphrased content with high accuracy.

Configuring the Django Framework and BERT Model for Local Deployment:

Assisted in setting up the Django application and integrating the BERT model locally, optimizing the environment for offline usage and improving loading times for faster processing.

Testing, Debugging, and Optimizing the System:

Performed extensive testing, identified bugs, and implemented fixes to ensure smooth functionality of the system. Focused on text processing issues, database interactions, and model performance optimization.

↓ User Interface Development and Feedback Integration:

Contributed to the design of a user-friendly interface for uploading projects, checking plagiarism, and displaying results. Iteratively improved the interface based on user feedback to enhance usability and user experience.

Analysis of Learned Knowledge and Applied Skills:

Integration of Machine Learning Models with Web Frameworks:

Learned and applied the process of integrating machine learning models, specifically for text analysis, with a web development framework (Django). This included model deployment, API integration, and seamless communication between the backend model and the web application.

♣ Proficiency in Python, Django, and BERT:

Gained extensive hands-on experience using Python and the Django framework for building robust web applications. Acquired practical knowledge of working with the BERT model for natural language processing tasks, particularly in embedding text data and conducting semantic similarity checks for plagiarism detection.

Enhanced System Design and Development Skills:

Learned how to design and develop a comprehensive AI-powered plagiarism detection system. This included implementing document processing pipelines, performing text similarity analysis, and building efficient data storage mechanisms for unique academic projects.

↓ Improved Problem-Solving and Debugging Techniques:

Applied analytical thinking and debugging skills to address challenges in model integration, performance optimization, and data handling. Gained experience in troubleshooting issues related to text analysis, model performance, and scalability.

Experience in Full-Stack Development:

Acquired practical skills in full-stack web development, involving both frontend and backend components of the plagiarism detection system. This included creating user-friendly interfaces, managing server-side logic, and ensuring secure interactions between users and the system.

A Profile of Skills and Competencies Gained/Acquired:

▶ Proficiency in Python and Django Web Development:

Gained advanced skills in Python programming and web application development using the Django framework. This includes creating and managing databases, implementing server-side logic, and building responsive, user-friendly web interfaces.

▶ Machine Learning Model Integration:

Acquired experience in integrating machine learning models for real-world applications, specifically for text analysis and plagiarism detection. Learned how to preprocess text data, fine-tune models, and evaluate their performance within a web-based system.

> Practical Experience in AI System Development:

Gained hands-on experience in developing an AI-powered plagiarism detection system, including document analysis, similarity detection, and result interpretation. Improved skills in handling large datasets, embedding text data, and optimizing model performance.

Database Management and Data Storage:

Developed competency in designing and managing relational databases, with a focus on securely storing unique project documents after plagiarism checks. Enhanced skills in SQL and Django's ORM for efficient data retrieval and storage.

> Teamwork and Communication Skills:

Improved collaboration skills through regular interactions with supervisors, team members, and users. Gained experience in presenting project progress, receiving feedback, and implementing changes based on input from stakeholders.

> Problem-Solving and Adaptability:

Strengthened problem-solving skills by addressing challenges related to system performance, data handling, and model deployment. Demonstrated adaptability in overcoming obstacles and adjusting to the resource limitations faced during the attachment period.

CHAPTER 3: INDUSTRIAL ATTACHMENT LIFE-EXPERIENCE

What I Learned:

During my industrial attachment, I gained valuable hands-on experience in applying AI and machine learning to solve real-world problems. Working on the anti-plagiarism system allowed me to understand the practical application of machine learning models like BERT in text analysis and similarity detection. I also learned about web development frameworks like Django, which was essential in building the system's user interface and backend. The attachment provided insights into integrating AI models into a functional web application, making the theoretical concepts I learned in class more tangible.

What I Didn't Learn:

While I gained significant exposure to AI and machine learning concepts, the project's scope did not delve deeply into advanced deep learning techniques or more complex neural networks, such as convolutional or recurrent neural networks. Additionally, I did not explore areas like model optimization or hyperparameter tuning to the depth I had hoped for, which would have added

another layer of learning. These are areas I would like to explore further in my future work or academic studies.

Relevance of the Experience to My Training:

This attachment was highly relevant to my academic training as it allowed me to bridge the gap between theoretical knowledge and practical application. The opportunity to apply concepts such as natural language processing, machine learning, and web development in a real-world project not only solidified my understanding but also enhanced my problem-solving skills. It gave me a comprehensive view of how AI systems are developed and deployed in a real-world context, making it a vital part of my learning journey.

CHAPTER 4: GENERAL CONCLUSION AND RECOMMENDATIONS

Summary:

This report outlines the work completed during my seven-week industrial attachment at RP Musanze College, where I was assigned to the IT department as an AI developer. My primary responsibility was to develop and implement an anti-plagiarism system for academic research projects. The system is designed to check uploaded student projects for plagiarism, ensuring that only unique and original work is accepted and stored in the database. By integrating AI-based text analysis, the system can accurately identify similarities between the uploaded documents and existing projects, helping to maintain academic integrity.

The anti-plagiarism system not only performs similarity checks but also serves as a comprehensive project management tool, automating traditional steps involved in the submission and evaluation of student projects. This includes features for project upload, automatic plagiarism detection, and secure storage of verified unique projects. The system aims to streamline the process of project review, enhance transparency, and provide a reliable platform for students and academic staff.

Overall, this industrial attachment provided a valuable opportunity to apply my academic knowledge in a practical setting, refine my skills in AI and software development, and contribute to a meaningful solution that addresses a key challenge in academic research. The report captures the scope of the project, the skills gained, the challenges faced, and recommendations for future improvements.

Conclusion and Recommendations:

Conclusion

My industrial attachment at RP Musanze College, within the IT department, was an enriching experience that allowed me to apply my academic knowledge to real-world scenarios. Working as an AI developer on the anti-plagiarism system for academic research projects provided me with valuable hands-on experience in software development and problem-solving. I was able to enhance my skills in programming, AI model integration, and system design, all while contributing to a meaningful project that addresses a significant issue in academic research.

Throughout the seven weeks of attachment, I encountered some challenges, particularly regarding the availability of a dedicated workspace, which occasionally impacted my ability to carry out tasks efficiently. However, these challenges taught me adaptability and the importance of resource management in a professional setting. The mentorship and support I received from the department staff were instrumental in helping me overcome obstacles and successfully complete the project tasks.

Overall, the attachment experience was highly beneficial and provided a strong foundation for my professional growth. It has given me a clearer perspective on my career goals and further motivated me to pursue advanced skills in AI and software development. I am grateful for the opportunity and look forward to applying the knowledge gained in future endeavors.

Recommendations

During my industrial attachment at RP Musanze College in the IT department, I had the opportunity to work as an AI developer on an anti-plagiarism system for academic research projects. While the experience was valuable and provided me with practical skills in software development, there were some challenges that affected my productivity. At times, I struggled to find a suitable and dedicated workspace, which made it difficult to focus on my daily activities and tasks.

I recommend that the college consider providing a designated area or workspace for interns, equipped with the necessary IT resources such as computers and stable internet access. This would create a more conducive environment for interns to work effectively on their projects. Additionally, establishing a structured mentorship program and clear communication channels would ensure that interns receive continuous guidance and support throughout their attachment period.

Lastly, enhancing the onboarding process and offering broader exposure to other IT activities would greatly benefit interns by helping them integrate quickly and gain a well-rounded learning experience. Implementing these improvements would not only enhance the internship program but also foster a more productive and supportive environment for future interns.

REFERENCES

10

https://www.grammarly.com/plagiarism-checker?utm_source=google&utm_medium=cpc&utm_campaign=19841729275&utm_c



ontent=652327824443&utm_term=plagiarism%20checker&target=&targetid=aud-1725360773014:kwd-

11839001&adgroup=146708454909&device=c&matchtype=b&placement=&network=g &extension=&clickid=Cj0KCQiA6Ou5BhCrARIsAPoTxrBJ12JTlO6Wf0eOD3wpGyze wWnJpN5nUQbXRY5ZTOzgI12oxP3P51QaAuoLEALw wcB&gad source=1&gclid=Cj0KCQiA6Ou5BhCrARIsAPoTxrBJ12JTlO6Wf0eOD3wpGyzewWnJpN5nUQbXRY5ZTOzgI12oxP3P51QaAuoLEALw wcB&gclsrc=aw.ds

https://huggingface.co/blog/bert-101

APPENDICES

• Appendix 2: Code Snippets

Index.html codes:

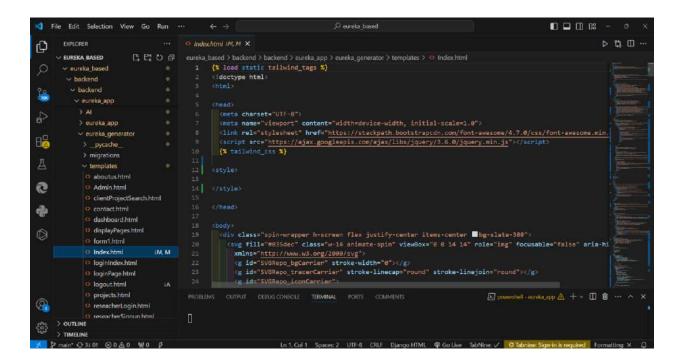


Figure 2

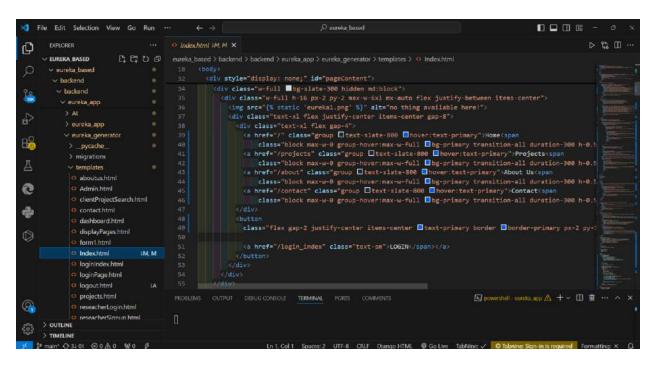


Figure 3

About Us codes:

12

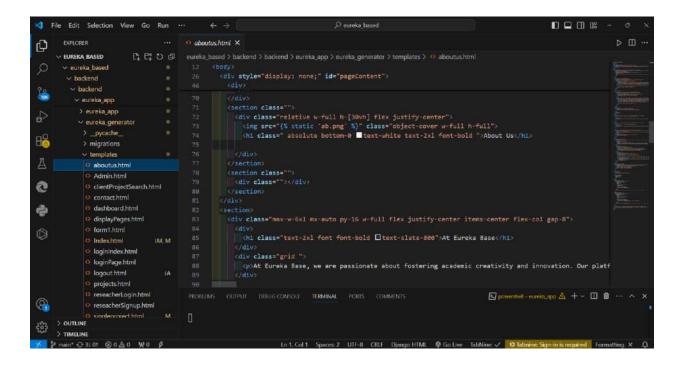


Figure 4

Admin and Users Pages Codes:

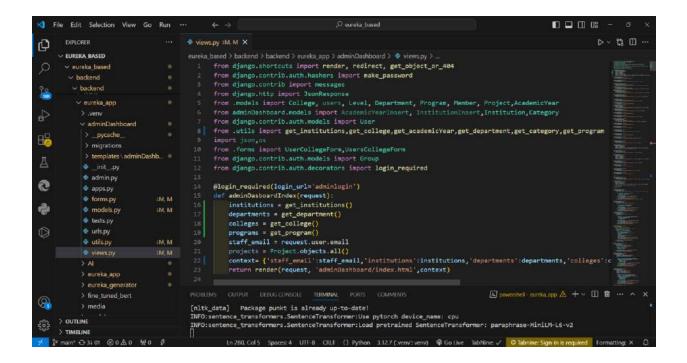


Figure 5

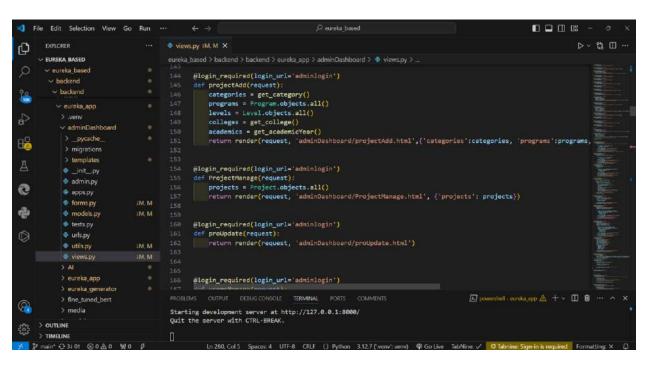


Figure 6

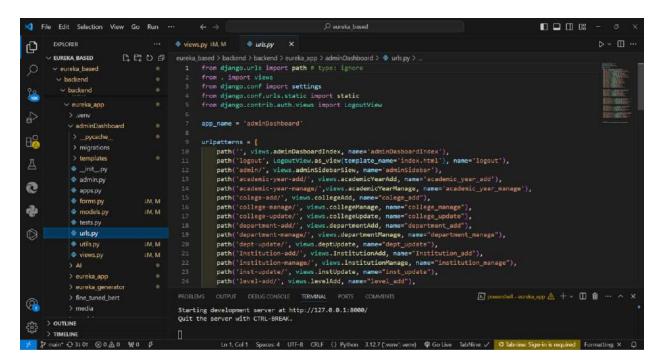


Figure 7

AI Training Codes:

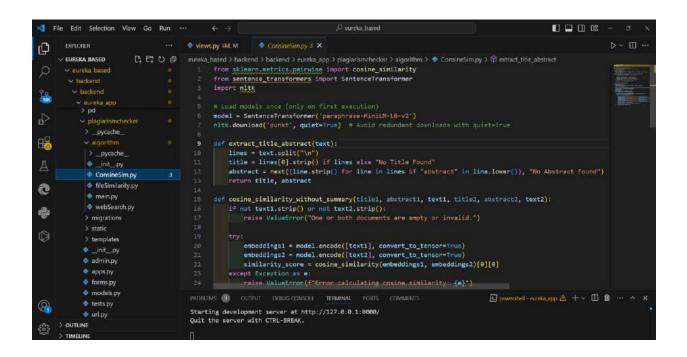
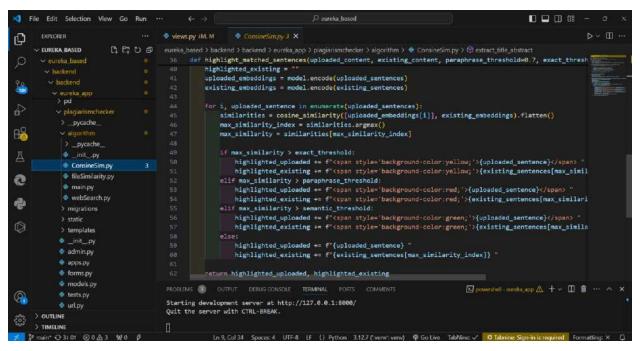


Figure 8



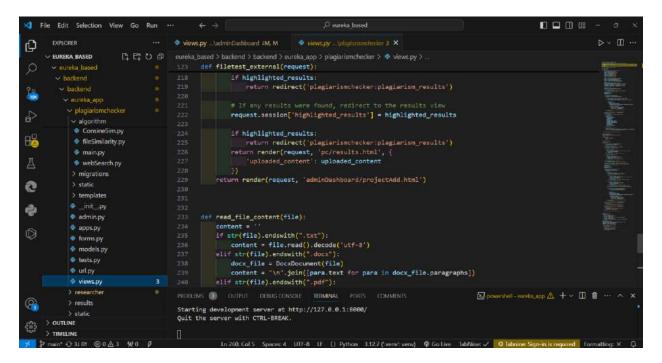


Figure 10

17

• Appendix 3: Test Results

```
[18/Nov/2024 15:11:31] "GET /static/eureka1.png HTTP/1.1" 200 2895
[18/Nov/2024 15:11:31] "GET /static/bg1.jpg HTTP/1.1" 200 866940
[18/Nov/2024 15:11:31] "GET /static/brain-organ.png HTTP/1.1" 200 630
Not Found: /js/jquery-ui.js
WARNING:django.request:Not Found: /js/jquery-ui.js
[18/Nov/2024 15:11:31] "GET /js/jquery-ui.js HTTP/1.1" 404 8011
Not Found: /js/jquery-ui.js
WARNING:django.request:Not Found: /js/jquery-ui.js
[18/Nov/2024 15:11:34] "GET /js/jquery-ui.js HTTP/1.1" 404 8011
Not Found: /favicon.ico
WARNING:django.request:Not Found: /favicon.ico
[18/Nov/2024 15:11:36] "GET /favicon.ico HTTP/1.1" 404 7999
[18/Nov/2024 15:11:38] "GET /projects HTTP/1.1" 301 0
D:\EurekaBase\eureka_based\eureka_based\backend\backend\eureka_app\eureka_generator\views.py:143: UnorderedObjectListWarning: Pa
gination may yield inconsistent results with an unordered object_list: <class 'eureka_generator.models.Project'> QuerySet.
  paginator = Paginator(projects, 6) # Show 6 projects per page
[18/Nov/2024 15:11:38] "GET /projects/ HTTP/1.1" 200 24439
[18/Nov/2024 15:11:38] "GET /static/css/dist/styles.css?v=1731935498 HTTP/1.1" 200 37900 [18/Nov/2024 15:11:38] "GET /static/js/jquery-ui.js HTTP/1.1" 200 106
first member is None
[18/Nov/2024 15:11:41] "GET /singleproject/2/ HTTP/1.1" 200 9630
[18/Nov/2024 15:11:41] "GET /static/css/output.css HTTP/1.1" 404 1856
[18/Nov/2024 15:11:41] "GET /static/css/dist/styles.css?v=1731935501 HTTP/1.1" 200 37900 [18/Nov/2024 15:11:46] "GET /about HTTP/1.1" 301 0
[18/Nov/2024 15:11:46] "GET /about/ HTTP/1.1" 200 14097

[18/Nov/2024 15:11:46] "GET /static/css/dist/styles.css?v=1731935506 HTTP/1.1" 200 37900

[18/Nov/2024 15:11:46] "GET /static/ab.png HTTP/1.1" 200 671857
[18/Nov/2024 15:11:49] "GET /contact HTTP/1.1" 301 0
[18/Nov/2024 15:11:49] "GET /contact/ HTTP/1.1" 200 14692
[18/Nov/2024 15:11:49] "GET /static/cureka2.png HTTP/1.1" 404 1847

[18/Nov/2024 15:11:49] "GET /static/css/dist/styles.css?v=1731935509 HTTP/1.1" 200 37900

[18/Nov/2024 15:11:52] "GET / HTTP/1.1" 200 24079
Not Found: /js/jquery-ui.js
               Ln 260, Col 5 Spaces: 4 UTF-8 LF () Python 3.12.7 (".venv": venv) 🗣 Go Live TabNine: 🗸 👲 Tabnine: Sign-in is required Formatting: X 🚨
```

Figure 11

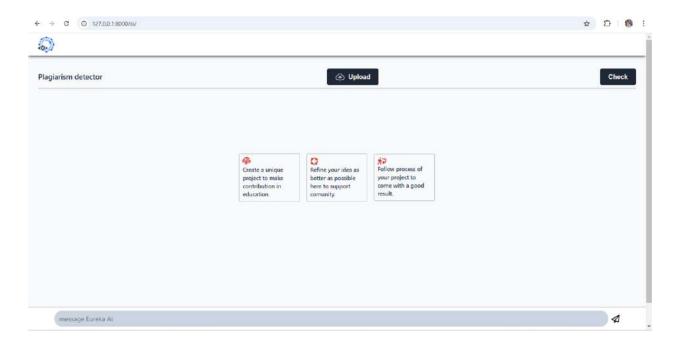


Figure 12

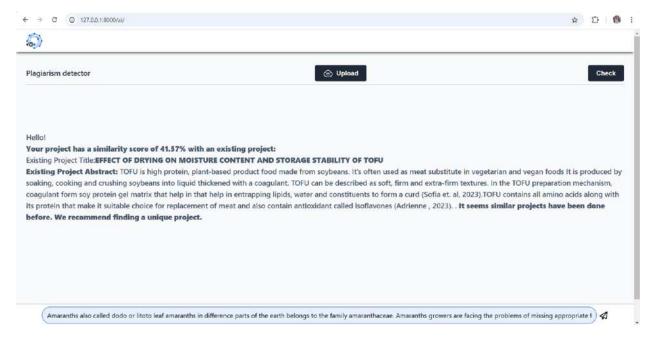
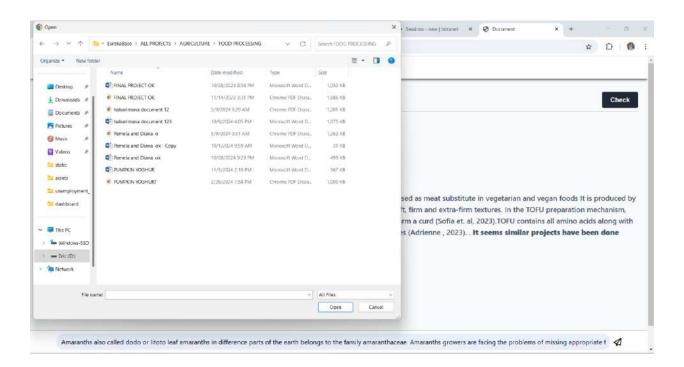
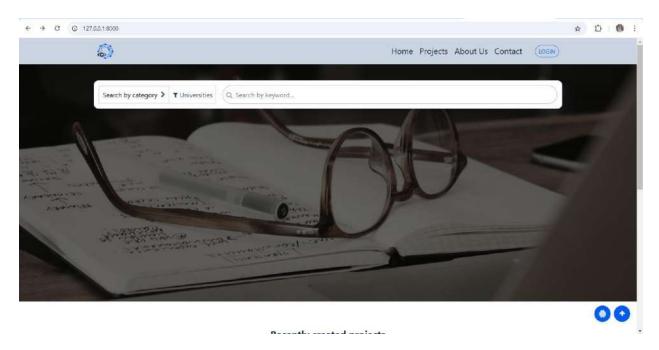


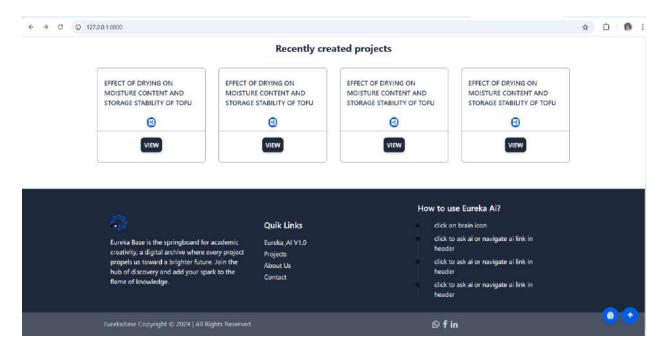
Figure 13



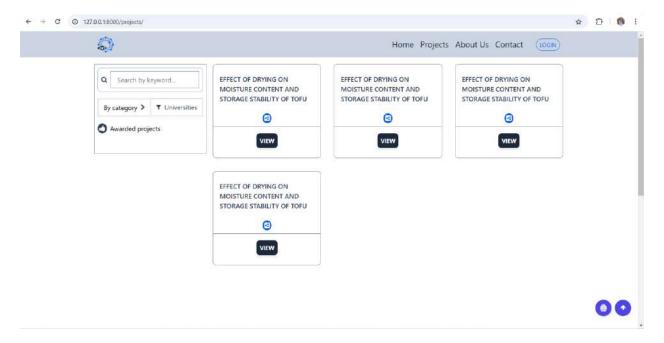
• Appendix 4: Screenshots of the system in use



Home Pagel Figure 15



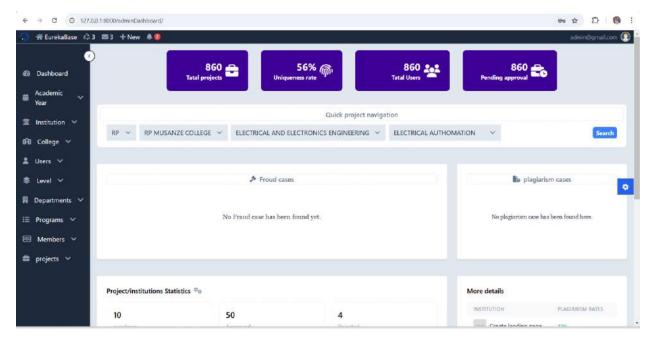
Home page2 Figure 16



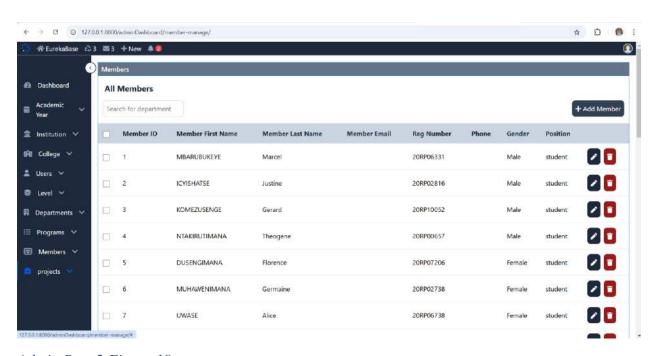
Project page Figure 17



Admin Dashboard



Admin page1 Figure 18



Admin Page2 Figure 19



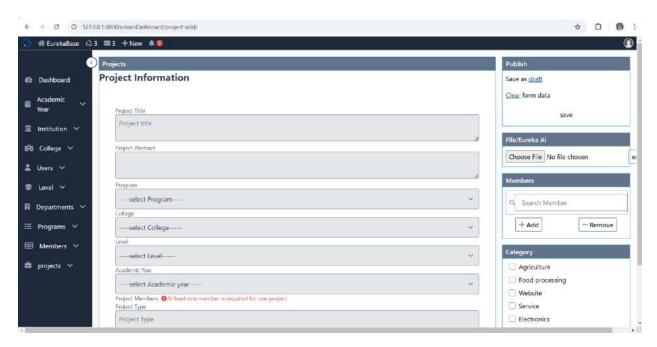


Figure 20 – Admin Saving project page

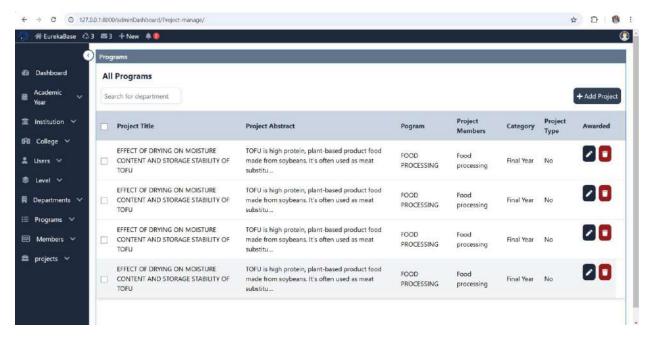


Figure 21 – Recorded Project with no plagiarism detected.

23

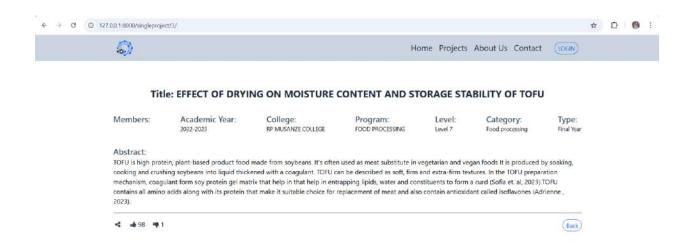


Figure 22 – Project overview page

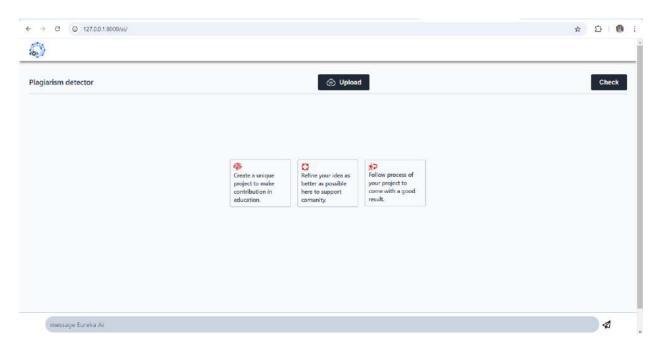


Figure 23 – Plagiarism checker Home page

Plagiarism Checker output

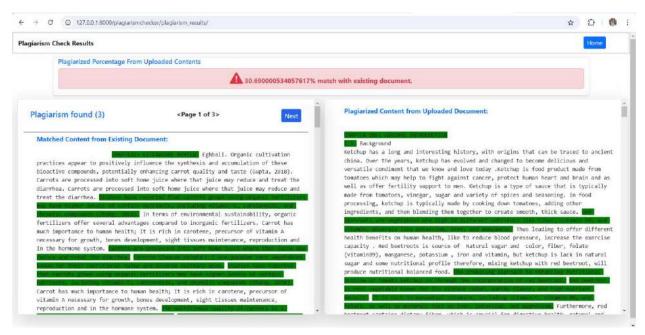


Figure 24 – output 1

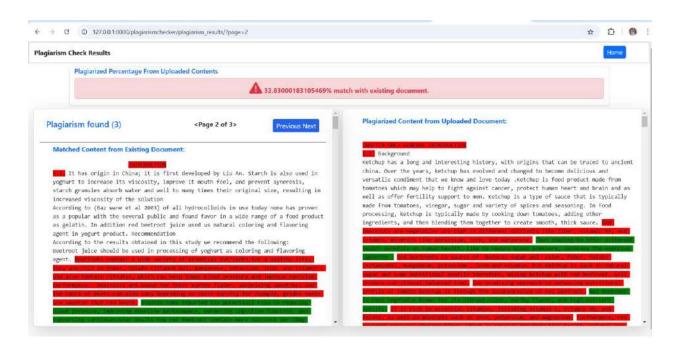


Figure 25 – output 2

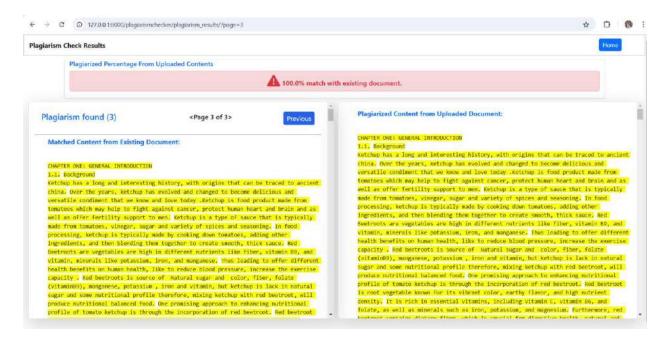


Figure 26 – output 3