CAPSTONE PROJECT

AI-DRIVEN PLAGIARISM FOR ASSIGNMENTS

Presented By:
1. Neha Sara Philip-Dayanand Sagar University-CSE Cyber Security

edunet foundation

OUTLINE

- **Problem Statement** (Should not include solution)
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT

The increasing ease of accessing online content and AI-generated text has led to a significant rise in plagiarism in academic assignments. Traditional plagiarism detection tools primarily compare textual similarity and often fail to detect intelligent paraphrasing or semantic modifications, especially when AI tools are used to disguise plagiarism.

PROPOSED SOLUTION

- AI-Driven Plagiarism Intelligence System that uses foundation models from IBM Watsonx (specifically, Granite series models) to semantically compare submitted assignments with historical data. The system leverages prompt engineering to analyze context, detect paraphrasing, and flag potential plagiarism cases.
- 1. Data Collection
- Collect past assignment data
- Accept new assignment input
- 2. Data Processing
- Clean and tokenize text
- Convert text to embeddings
- 3. Machine Learning Algorithm
- Use Granite model for comparison
- Classify as original/paraphrased/copied
- 4. Deployment
- Deploy in Watsonx.ai Studio
- Store output as JSON/CSV
- 5. Evaluation
- Check accuracy on known cases
- Review model explanations

edunet foundation

SYSTEM APPROACH

1.System Requirements

IBM Cloud account (Free Lite plan)Stable internet for Watsonx Studio

Access Browser: Chrome/Edge with Jupyter support

4GB+ RAM recommended for local preprocessing

2. Library Requirements

openai / ibm-watsonx SDK (for Granite API)

pandas – data handling (CSV/JSON)

nltk or re – basic text preprocessing

json – reading/writing output

matplotlib (optional) – result visualization

ALGORITHM & DEPLOYMENT

- **Algorithm Selection:**
  - Chose Granite Foundation Model from IBM Watsonx
  - Suitable for semantic understanding and natural language comparison
- **Data Input:**
  - New assignment uploaded by user
  - Historical assignments fetched for comparison
- **Training Process:**
  - No custom training (used pre-trained model)
  - Custom prompts used to adapt model to plagiarism detection
- **Prediction Process:**
  - Model compares new vs. old assignments using prompt
  - Outputs result as Original, Paraphrased, or Copied

RESULT

```
/opt/conda/envs/Python-RT24.1/lib/python3.11/site-packages/ibm_watsonx_ai/foundation_models/utils/utils.py:436: LifecycleWarning: Model 'ibm/granite-13b-instruct-v2' is in deprecated state from 2025-06-18 until 2025-10-15. IDs of alternative models: ibm/granite-3-3-8b-instruct. Further details: https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/fm-model-lifecycle.html?context=wx&audience=wdp
  warn(model_state_warning, category=LifecycleWarning)
{'model_id': 'ibm/granite-13b-instruct-v2', 'created_at': '2025-07-31T04:25:35.853Z', 'results': [{'generated_text': 'The student copied the code from ChatGPT.', 'generated_token_count': 11, 'input_token_count': 15, 'stop_reason': 'eos_token'}], 'system': {'warnings': [{'message': "Model 'ibm/granite-13b-instruct-v2' is in deprecated state from 2025-06-18. It will be in withdrawn state from 2025-10-15. IDs of alternative models: ibm/granite-3-3-8b-instruct.", 'id': 'deprecation_warning', 'more_info': 'https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/fm-model-lifecycle.html?context=wx&audience=wdp'}, {'message': "The value of 'parameters.max_new_tokens' for this model was set to value 20", 'id': 'unspecified_max_new_tokens', 'additional_properties': {'limit': 0, 'new_value': 20, 'parameter': 'parameters.max_new_tokens', 'value': 0}}, {'message': "This API is legacy. Please consider using '/ml/v1/text/chat' instead.", 'id': 'api_legacy'}]}}
```

```
/opt/conda/envs/Python-RT24.1/lib/python3.11/site-packages/ibm_watsonx_ai/foundation_models/utils/utils.py:436: LifecycleWarning: Model 'ibm/granite-13b-instruct-v2' is in deprecated state from 2025-06-18 until 2025-10-15. IDs of alternative models: ibm/granite-3-3-8b-instruct. Further details: https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/fm-model-lifecycle.html?context=wx&audience=wdp
  warn(model_state_warning, category=LifecycleWarning)
{'model_id': 'ibm/granite-13b-instruct-v2', 'created_at': '2025-07-31T04:33:02.596Z', 'results': [{'generated_text': 'Yes. The second text explains why recursion is used in computing.', 'generated_token_count': 14, 'input_token_count': 73, 'stop_reason': 'eos_token'}], 'system': {'warnings': [{'message': "Model 'ibm/granite-13b-instruct-v2' is in deprecated state from 2025-06-18. It will be in withdrawn state from 2025-10-15. IDs of alternative models: ibm/granite-3-3-8b-instruct.", 'id': 'deprecation_warning', 'more_info': 'https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/fm-model-lifecycle.html?context=wx&audience=wdp'}, {'message': "The value of 'parameters.max_new_tokens' for this model was set to value 20", 'id': 'unspecified_max_new_tokens', 'additional_properties': {'limit': 0, 'new_value': 20, 'parameter': 'parameters.max_new_tokens', 'value': 0}}, {'message': "This API is legacy. Please consider using '/ml/v1/text/chat' instead.", 'id': 'api_legacy'}]}}
```

```
/opt/conda/envs/Python-RT24.1/lib/python3.11/site-packages/ibm_watsonx_ai/foundation_models/utils/utils.py:436: LifecycleWarning: Model 'ibm/granite-13b-instruct-v2' is in deprecated state from 2025-06-18 until 2025-10-15. IDs of alternative models: ibm/granite-3-3-8b-instruct. Further details: https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/fm-model-lifecycle.html?context=wx&audience=wdp
  warn(model_state_warning, category=LifecycleWarning)
Comparing with historical submissions...

Comparison with Past Submission 1:
Yes

Comparison with Past Submission 2:
Yes

Comparison with Past Submission 3:
Yes

Instructor Feedback:
 This is suspicious because it is not a common programming concept.

Final Verdict:
The assignment is flagged as suspicious/plagiarized.
```

edunet foundation

CONCLUSION

The proposed system successfully demonstrates the use of AI foundation models in plagiarism detection by going beyond exact matches and identifying paraphrased or semantically similar content. It can be a powerful tool for academic institutions aiming to uphold originality and integrity in submissions.

FUTURE SCOPE

1. Integrate with college LMS like Moodle or Google Classroom
2. Add support for multi-language plagiarism detection
3. Introduce visual dashboard for teachers with statistics
4. Include plagiarism severity scoring
5. Add voice-to-text plagiarism detection (audio to text comparison)
6. Train a custom model using internal institutional data for higher accuracy

REFERENCES

List of Resources Used:

IBM Watsonx.ai Documentation

https://www.ibm.com/products/watsonx

Granite Foundation Model (IBM) Overview

https://www.ibm.com/blog/foundation-models

Python official documentation:

https://docs.python.org

GitHub for code sharing and version control Articles on semantic plagiarism detection using LLMs

IBM CERTIFICATIONS

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Neha sara Philip

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

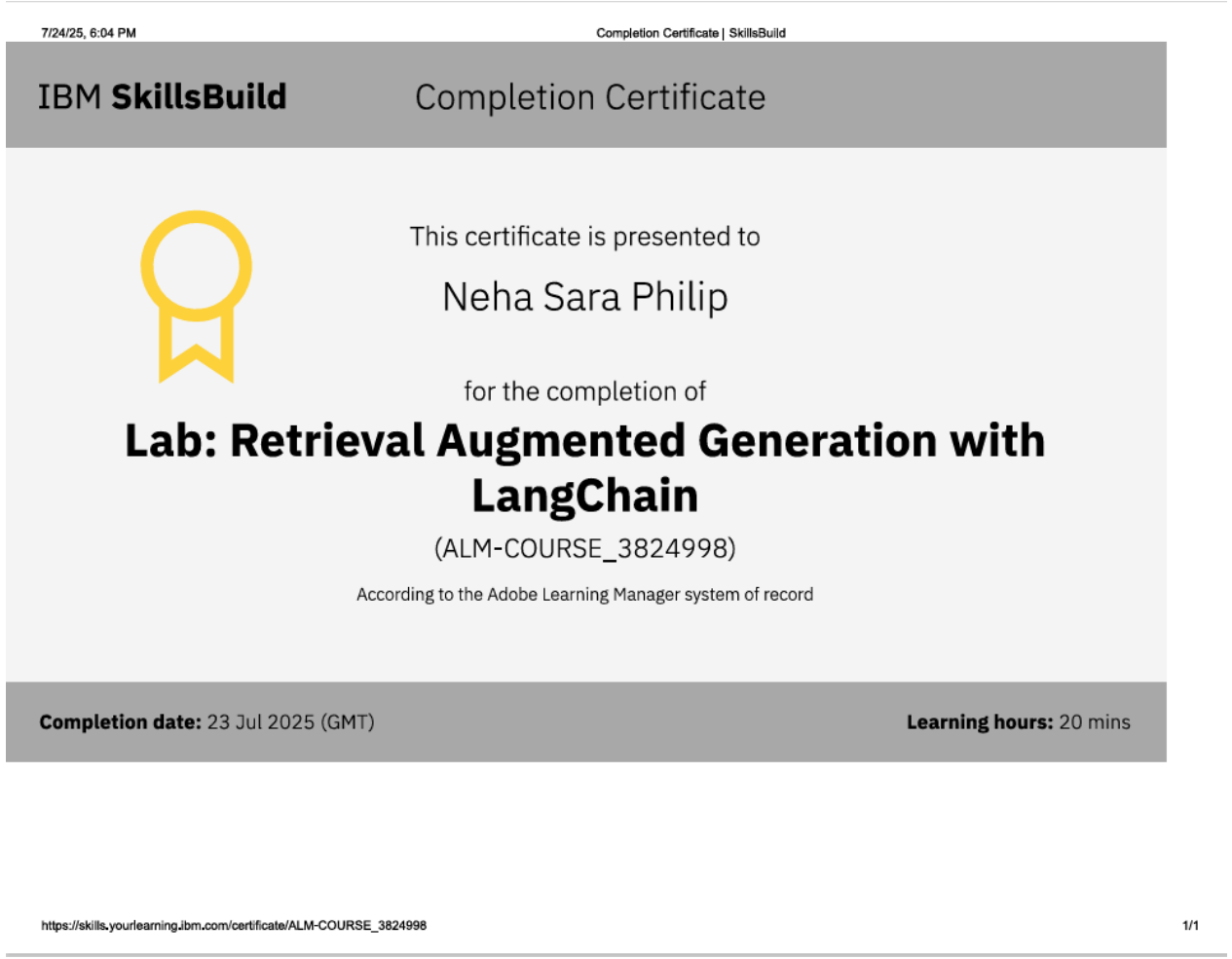
Issued on: Jul 17, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/cbc89bb8-94ff-4609-834a-4181ba7d5087

IBM

IBM CERTIFICATIONS



7/24/25, 6:04 PM Completion Certificate | SkillsBuild

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Neha Sara Philip

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 23 Jul 2025 (GMT) **Learning hours:** 20 mins

https://skills.yourlearning.ibm.com/certificate/ALM-COURSE_3824998 1/1

THANK YOU

edunet foundation