CAPSTONE PROJECT

AI-DRIVEN PLAGIARISM INTELLIGENCE FOR ASSIGNMENTS

Presented By:
Divaesh Nandaa - Anand Institute of Higher Technology - CSE

OUTLINE

- **Problem Statement**

- **Proposed System/Solution**

- **System Development Approach**

- **Algorithm & Deployment**

- **Result**

- **Conclusion**

- **Future Scope**

- **References**

PROBLEM STATEMENT

Academic institutions struggle to detect sophisticated forms of plagiarism in assignments, especially when text is paraphrased, restructured, or generated using AI tools.

Traditional plagiarism detectors rely heavily on surface-level similarity and fail to account for instructor-specific expectations, historical grading patterns, or writing context.

edunet
foundation

PROPOSED SOLUTION

- This system uses IBM Watsonx.ai and foundation models to detect subtle and AI-generated plagiarism with greater contextual awareness. It adapts to individual instructors' expectations and student writing styles by learning from historical submissions and feedback, offering tailored risk assessments.

Core Components:

- **Data Ingestion:**
  Aggregates student assignments, instructor feedback, grading rubrics, and academic datasets to build stylistic baselines.

- **Preprocessing & Style Modeling:**
  Text is cleaned and vectorized into unique author embeddings, capturing tone, structure, and linguistic patterns.

- **AI Analysis via Watsonx.ai:**
  Utilizes IBM Granite models and prompt-based NLP to:

  - Assess likelihood of AI-authored content

  - Measure stylistic divergence from known student writing or instructor expectations

- **Platform & Interface:**
  Deployed on IBM Cloud Lite with a secure, web-based dashboard for educators to upload and evaluate submissions easily.

- **Performance Metrics:**
  Evaluation based on precision, recall, F1-score, and semantic similarity thresholds to ensure accurate and fair assessments.

edunet
foundation

SYSTEM APPROACH

- *AI Model - IBM Granite (Foundation Model)*

- *Cloud Infrastructure - IBM Cloud Lite*

- *Prompt Engineering - Watsonx.ai Prompt Lab*

- *Input - Student Assignments, Instructor Style*

- *Output - Similarity (0-100)%, AI Probability (0-100)%, Risk Levels (0-100)%*

ALGORITHM & DEPLOYMENT

Model:

Uses IBM Watsonx.ai and Granite LLMs with prompt-based analysis to detect plagiarism, AI authorship, and writing inconsistencies.

Inputs:

- Current student assignment

- Previous submissions by the student

- Instructor rubrics, tone, and feedback

- External academic and plagiarism databases

- AI-confidence scores (model-generated)

Process:

Prompts guide the model to:

- Check for semantic similarity with known sources

- Estimate AI-generation likelihood

- Compare writing style to student's history

Outputs:

- **Similarity Score** (% match)

- **AI Probability** (likelihood of AI use)

- **Style Deviation** (from prior work)
  Visualized in a web dashboard or downloadable format.

edunet
foundation

RESULT

RESULT

CONCLUSION

The AI-driven plagiarism detection system is designed to meet the evolving demands of academic integrity by utilizing IBM Watsonx.ai and Granite Foundation Models. It goes beyond traditional detection tools by identifying subtle forms of plagiarism—including paraphrased and AI-generated content—through advanced prompt-based semantic analysis and contextual writing style evaluation.

What sets this system apart is its ability to adapt to individual instructor expectations, assess deviations in writing style, and estimate the likelihood of AI-generated content. This results in more accurate, personalized, and actionable feedback for educators.

- Development Challenges Faced:

- Obtaining ethically sourced and diverse historical assignment data

- Ensuring data privacy while enabling personalized writing style modeling

- Keeping pace with the fast-evolving capabilities of generative AI tools

- Future Enhancements:

- Incorporating contextual data from more institutions and instructor profiles

- Enhancing the explainability of results to improve trust and transparency

- Enabling real-time adaptive feedback mechanisms to refine prompt strategies and boost detection accuracy

edunet
foundation

FUTURE SCOPE

Building on the foundation of Watsonx.ai and Granite models, several enhancements can be introduced to improve scalability, accuracy, and user experience of the plagiarism detection system:

- Broader Data Integration: Connect with external academic databases, citation networks, and AI-content detection APIs to improve semantic match coverage and enhance plagiarism insights.

- Prompt Optimization: Refine and expand prompt structures in Watsonx Prompt Lab to improve detection granularity and adapt to more academic writing styles without retraining models.

- Multi-Institution Deployment: Scale the system across universities, departments, and online education platforms (e.g., Coursera, edX) for global academic integrity monitoring.

- Instructor-Customized Feedback: Generate personalized, context-aware feedback for students, enabling the system to serve as both a detection tool and a learning assistant.

- Edge Deployment Options: Use lightweight Watsonx-compatible inference layers for local deployment within institutional networks to ensure faster response times and better privacy control.

- Real-Time Writing Assistance: Develop plugins for popular editors (Google Docs, Word) to provide live plagiarism risk scores as students write, encouraging responsible authorship proactively.

REFERENCES

The following sources and research materials were instrumental in shaping the development and design of this AI-driven plagiarism detection system using IBM Watsonx.ai and Granite Foundation Models:

- IBM Research. (2022). AI and Education: Enabling Student Success with Integrity.

  IBM AI Research Whitepaper.

  https://www.ibm.com/research.

- Turnitin (2023). AI Writing and Plagiarism: Insights from Turnitin's Detection Research.

  https://www.turnitin.com

IBM CERTIFICATIONS

In recognition of the commitment to achieve
professional excellence

Getting Started with
Artificial Intelligence
IBM SkillsBuild
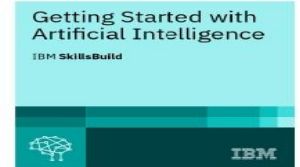
Divaesh Nandaa

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 16, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/1640b8d6-6884-4296-aa5a-ade368705909

IBM

edunet
foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

Divaesh Nandaa

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 20, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/9b024c45-efb0-4790-8ae7-c252e9a74337

IBM

eaunet
foundation

IBM CERTIFICATIONS

IBM **SkillsBuild**

Completion Certificate

This certificate is presented to

Divaesh Nandaa

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 24 Jul 2025 (GMT)   **Learning hours:** 20 mins

edunet
foundation

THANK YOU