**INTRODUCTION**

Analyzing student performance on an e-learning platform involves leveraging advanced tools, including a pretrained machine-generated text detection model. Key metrics include response quality, time taken to respond, and engagement with professors’ instructions. Additionally, the analysis assesses adherence to deadlines, the ability to answer standardized questions consistently, and overall interaction patterns. This comprehensive approach provides a data-driven understanding of student behavior, offering valuable insights for improving learning outcomes.

**ATTRIBUTES**

1. id
2. userfullname
3. totalPosts
4. activeDays
5. total\_replies\_to\_prof
6. total\_characters
7. total\_words
8. avg\_words
9. unique\_interactions
10. unique\_discussions
11. engagement\_rate
12. normanlized\_score
13. deadline\_exceeded(Quasi exam I)
14. deadline\_exceeded(Quasi exam II)
15. deadline\_exceeded(Quasi exam III)
16. Pattern\_followed(quasi exam i)
17. avg\_AI\_involvedMsg\_score
18. User ID
19. Full Name
20. Total number of posts
21. Number of days that student actively participated in discussions
22. Number of replies to professor's post
23. Total chars of all posts of the student
24. Total words of all posts of the student
25. Avarage words per post
26. the number of unique interactions with other students for each user
27. the number of unique discussions for each user
28. The attribute measures student engagement by dividing their replies by the professor's total posts, showing how actively they participate in the professor's discussions.
29. The used query normalizes total replies and average reply time for all students, assigning a higher weight (70%) to replies and a lower weight (30%) to reply time. This weighting emphasizes the importance of engagement (quantity of replies) over responsiveness (speed of replies) when calculating an overall performance score. The normalized scores ensure fair comparisons regardless of differing scales for the two metrics.
30. Posts created after given deadline in quasi exam I
31. Posts created after given deadline in quasi exam III
32. Posts created after given deadline in quasi exam III
33. The number of posts where a student followed a specific pattern (with a maximum of 2 and a minimum of 0)
34. The average AI involvement score is calculated using the "roberta-base-openai-detector" model, which detects AI-generated text. It assigns a score for each response (scaled 1-10 the more the high probability of AI involvement), and averaging these scores indicates the overall prevalence of AI-generated content.

**STEPS**

**Data preparation**

1. Converting excel file into CSV file to import into database
2. Converting time format into database compatible format (iso 8859-1)
3. Pulling data into database

**Data extraction**

1. Writing queries. Each and every query used for getting relevant information is attached.

**AI involvement analysis**

1. Prepare the Environment – getting all important python libraries. (Source code attached)
2. Load and Preprocess Data.
3. Function to clean text (Remove punctuation, Normalize whitespace)
4. Analyze Text Complexity (Calculate readability metrics)
5. Predict AI probability
6. Combine Attributes into a Unified Score
7. Scale to a 1-10 rating
8. Pull the result into the database

**Creation of OAM**

1. Combine and rank all the gathered data into single OAM
2. Give a direction to all attributes 0, 1

**COCO Y0**

1. Using COCO AI to determine the performance of all students

**Writing Documents**