CAT2

1) Which type of machine learning algorithm would be most suitable for this task?

Explain your reasoning.

data-driven models.

Data-driven models excel at recognizing patterns and relationships within large datasets. This is particularly useful in educational settings where student behavior, academic performance, and other relevant factors can be complex and interconnected.

2) What features from the student data would be most relevant for predicting enrollment and graduation success?

Demographic Information

Academic Performance

3) How can you protect the privacy of student data while still using it to develop predictive models?

Remove or encrypt personally identifiable information (PII) such as names, addresses, and social security numbers to ensure that individual identities cannot be easily discerned.

Aggregate data to a level where individuals cannot be identified. For example, use group averages or cohorts instead of individual student records.

Limit Access**:**

Restrict access to the data to only those individuals who need it for model development. Implement strong access controls and authentication mechanisms to prevent unauthorized access.

4) How can you communicate the results of your model to educational institutions in a way that is actionable and informative?

Compare the model results with relevant benchmarks or historical data to provide context. This helps institutions understand how their students compare to broader trends.

Clearly identify the groups of students identified as high-risk and provide information on the factors contributing to their risk status. This enables targeted interventions