Your Name: Vani Singh

MSDS 603 MLOps Assignment 1 – Part 1 (2.5%)
Requirements Gathering

In this assignment, you will gather the requirements for building a specific AI/ML-powered product. You will need to identify the business and technical requirements, assess potential risks, propose mitigation strategies, and outline the high-level components needed for successful implementation of the product. You **will not** need to actually build the product.

Learning Objectives

- Apply MLOps principles to a real-world product scenario
- Practice requirements gathering and analysis for ML systems
- Identify potential risks in ML systems and develop mitigation strategies
- Understand the core components required in an ML product pipeline

Scenario

An EdTech company is developing a personalized learning platform for K-12 students. The platform will use machine learning and AI to analyze student performance data from standardized tests and ongoing assessments within the platform to create customized learning pathways for each student. The system should adapt in real-time to student progress, identifying knowledge gaps, recommending appropriate learning activities, and adjusting difficulty levels to maximize learning outcomes while maintaining student engagement. The platform must eventually work across various subjects, but for now we will focus only on ***reading comprehension***. Ideally, it should accommodate different learning styles, comply with educational privacy regulations (like FERPA), and provide actionable insights to teachers and parents through intuitive dashboards.

Requirements

This assignment is done in **two parts**. Complete Part One below in class, and turn in before leaving class. Complete Part Two at home and turn in to Canvas. If you did not attend class for Part One, you must accept a zero grade for this assignment since Part Two depends on your answers to Part One.

Part One

In this part, take 30 minutes to answer the below questions in the space provided. No use of technology is allowed, just use your brain. No working with a partner, you must do this on your

own. There are no ***right*** or ***wrong*** answers here, but lack of effort will affect your grade. Please write as clearly and neatly as possible. Do not forget to write your name at the top of this page.

Question 1: Define two goals for this project.

1. Be able to accurately measure student performance and identify knowledge gaps to improve performance
2. Ensure the student stays engaged.

Question 2: For each goal from Question 1; define a metric to measure success of that goal.

1. Average improvement per student at the end of a semester
   ↳ $\frac{1}{n} \Sigma$ (grade at the end of the sem - grade at the beginning)
2. ~~Averega~~ Average time of active use per student per day in minutes.

Question 3: Identify at least three sources of data that will be needed.

1. Personal data of the student's performance ⇒ can be collected through evaluation activities and manual entry of grade
2. National data on student performance to ensure the student is performing well in comparison to their peers.
3. Educational data and materials from schools to ensure consistency in the curriculum being taught.

Question 4: Identify two risks associated with this product and the potential impact of each risk.

1. The model misclassifies the student's performance ⇒ if it recommends exercises that are too hard or too easy as a result, this will decrease student engagement.

2. The model recommends the wrong learning activities ⇒ the student's performance, despite maybe improving in the program, will not improve in school.

Question 5: For each risk identified in Question 4; propose a strategy to mitigate the risk.

1. Have multiple 'performance checks' to reduce the likelyhood of misclassification.

2. curriculum should be the top most prioritized feature when suggesting content for the student.

Question 6: Outline the major architectural components needed for this product.

1. Data storage

2. Model for evaluation

3. Model for recommendation

4. UI