CH5019 - Mathematical Foundations of Data Science
Course Project

An instructor wants to grade answers to descriptive questions automatically. The instructor has a template best answer that she/he has developed and wants to use the same to grade descriptive answers of students. The comparison results between the template and the student provided answer could be categorical (right vs wrong) or continuous (75% similarity and so on). You are expected to develop an AI algorithm that can do this comparison automatically. Please develop an algorithm through any appropriate concept and demonstrate the results on ten test cases. You can use any training approach and test it on any 10 test cases that you feel are appropriate. Your algorithm should take two paragraphs, one correct answer and one student provided answer and return a result. Since this is a rather open-ended problem, the solutions will be largely evaluated based on the creativity associated with problem formulation, solution approach (including feature engineering), and the achieved results. Please remember that there is no single correct method. Please use existing AI/ML libraries as much as possible.

Instructions:

- Groups up to a maximum of five members can be made.
- Code that can be run by the TAs on our own test cases should be submitted.
- A detailed report on the method used, training approach and test cases should be submitted. The maximum page limit for this is 10.
- A contribution statement that clearly articulates the contribution of each group member is mandatory at the end of the report. Reports without contribution statement will not be graded.