**University of Sharjah**

**College of Computing and Informatics**

Department of Computer Science

**1511330: Intro to Artificial Intelligence**

**Fall 2024 – 2025**

Due date: 20/11/2024 11:59PM

**Group Project**

****

**Integrity Statement: Read carefully!**

Strict directives have been given to the assistants who will be marking this homework to follow the University’s policy on cheating. Any student caught **cheating** **in any form** will be given zero (0) the first time. (We will not get into the issue of who copied from whom!!!). If such act reoccurs, the case will be sent to the University Student Discipline Committee.

**Group Project overview:**

In this group project, you are asked to explore the algorithms of Artificial Intelligence to the problem of your choice. You can choose to study any problem, giving special consideration to the unique properties of the problem domain, and testing two algorithms or more on it.

**Instructions:**

**1. Choose a Problem:** (weightage = 5%)

Visit the UCI Machine Learning Repository (<https://archive.ics.uci.edu/ml/index.php>) or <https://www.kaggle.com/datasets/> and select a dataset that you find interesting. Describe the dataset briefly, including the number of features, the class labels, and any challenges or unique properties of the dataset.

**2. Related works** (weightage = 5%)

You are required to search for FIVE (5) scholarly articles related to the topic(s) you have chosen using <https://scholar.google.com/> . Summarize each article and provide key insights.

**3. Data Preprocessing:** (weightage = 15%)

Clean and preprocess the dataset as necessary. This may include handling missing data, encoding categorical features, and normalizing/standardizing numerical features.

**4. Select Machine Learning Algorithms:** (weightage = 15%)

Choose at least three different algorithms for this task as these three algorithms will be compared with each other. Explain why you chose these algorithms and how they are suitable for your selected problem.

**5. Model Development:** (weightage = 40%)

Implement the selected algorithms using the preprocessed dataset. Clearly document the process, including hyperparameter tuning and cross-validation methods used for training.

**6. Model Evaluation:** (weightage = 10%)

Evaluate the performance of each algorithm using appropriate evaluation metrics such as accuracy, precision, recall, F1-score, and ROC curves, and standard deviation.

**7. Analysis:** (weightage = 10%)

Provide a detailed analysis of the results. Explain why certain algorithms performed better or worse on this specific problem. Discuss the implications of the unique properties of your chosen dataset on model performance. Moreover, elaborate the outcome of standard deviation result.

**8. Conclusion and Recommendations:**

Conclude the assignment by summarizing your findings and providing recommendations on which algorithm is the most suitable for the selected problem and why.

**Learning Outcomes:**

* Demonstrate ability to use inductive learning techniques to solve problems.

**Software**: Python programming

**Deliverables:**

Submission Requirements

**All submission is on the Blackboard.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Student declaration:** | | | | | |
| ***I declare that:***   * *I understand what is meant by plagiarism* * *The implication of plagiarism has been explained to us by our lecturer* * *This project is all my work and I have acknowledged any use of the published or unpublished works*   *of other people.* | | | | | |
| **Group:** | | | | | |
| **Names of Candidate** | | | | | |
| No. | **Student Name** | **Student ID** | **Contribution** | **Overall** | **Score** |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  |  |