University of Central Punjab

Faculty of Information Technology

Artificial Intelligence – Lab (F6)
Spring 2024

Instructor: Anaba Shafiq
Code: CSAL3241
Program: CS

Time Allowed: 100 Minutes **Total Marks:** 80

Instructions:

- Time to attempt the exam is 90 minutes. Last 10 minutes are for submission.
- Understanding of a question is part of exam. Don't ask invigilator for any explanation.
- Plagiarism will be penalized.
- Name your solution as "<your roll number>_mid_<section>", zip and submit.
- Implementation of both tasks in a single notebook is recommended.

Question 1: Genetic Algorithm **35 Marks**

You are tasked with optimizing the production scheduling for a manufacturing plant that produces 5 different products on 3 different machines. The goal is to minimize production time while ensuring each product's demand is met according to the following requirements:

- **Product 1**: Requires 4 hours of production time per unit.
- **Product 2**: Requires 5 hours of production time per unit.
- **Product 3**: Requires 3 hours of production time per unit.
- **Product 4**: Requires 6 hours of production time per unit.
- **Product 5**: Requires 2 hours of production time per unit.

The machines have different capacities and initial scheduling:

- **Machine A**: Currently assigned to Products 1, 2, and 3.
- **Machine B**: Currently assigned to Products 4 and 5.
- **Machine C**: Currently not assigned to any products.

Please outline the steps of a genetic algorithm approach to optimize this production scheduling, including details of the fitness function:

a) **Fitness Function**: Define the fitness function used to evaluate each production schedule. Include how you calculate the total production time and any penalties for failing to meet product demands.

b) **Parent Selection**: Describe how you would select parents for crossover based on their fitness scores.

c) **Crossover**: Explain the process of crossover (e.g., using a swap-based crossover) to generate new production schedules.

d) **Mutation**: Outline how you would introduce mutations (e.g., reassigning products to different machines) to explore new solutions.

e) **Termination**: Specify the termination criteria for the algorithm, such as achieving a certain level of efficiency improvement or after a fixed number of generations. Describe how you would determine the optimal production schedule.

Question 2: K-Nearest Neighborhood **45 Marks**

Wine Quality Data Set:

This data set contains records related to red and white variants of the Portuguese Vinho Verde wine. It contains information from 1599 red wine samples and 4898 white wine samples. Input variables in the data set consist of the type of wine (either red or white wine) and metrics from objective tests (e.g., acidity levels, PH values, ABV, etc.), while the target/output variable is a numerical score based on sensory data—median of at least 3 evaluations made by wine experts. Each expert graded the wine quality between 0 (very bad) and 10 (very excellent).

Input variables:

1. **type of wine:** type of wine (categorical: 'red', 'white')
2. **fixed acidity:** The acids that naturally occur in the grapes used to ferment the wine and carry over into the wine.
3. **volatile acidity:** Acids that evaporate at low temperatures—mainly acetic acid which can lead to an unpleasant, vinegar-like taste at very high levels.
4. **citric acid:** Citric acid is used as an acid supplement which boosts the acidity of the wine.
5. **residual sugar:** The amount of sugar remaining after fermentation stops.
6. **chlorides:** The amount of chloride salts (sodium chloride) present in the wine.
7. **free sulfur dioxide:** The free form of SO2 exists in equilibrium between molecular SO2 (as a dissolved gas) and bisulfite ion; it prevents microbial growth and the oxidation of wine.
8. **total sulfur dioxide:** The amount of free and bound forms of S02.
9. **density:** The density of wine juice depending on the percent alcohol and sugar content.
10. **pH:** A measure of the acidity of wine; most wines are between 3-4 on the pH scale.
11. **sulphates:** Amount of potassium sulphate as a wine additive which can contribute to sulfur dioxide gas (S02) levels.
12. **alcohol:** How much alcohol is contained in a given volume of wine (ABV).

Output variable:

- **quality:** score between 0 (very bad) and 10 (very excellent) by wine experts

The dataset is attached in the submission folder.

You have to perform the following tasks on the given dataset.

1. Read the data and explore it.
2. Separate labels from the data
3. Implement KNN
4. Perform prediction and calculate accuracy for k= 3, 5, 7, and 11.
5. Evaluate your model using different evaluation methods.

Note: Not allowed to use built-in ML Algorithm.