National University of Computer and Emerging Sciences, Lahore Campus

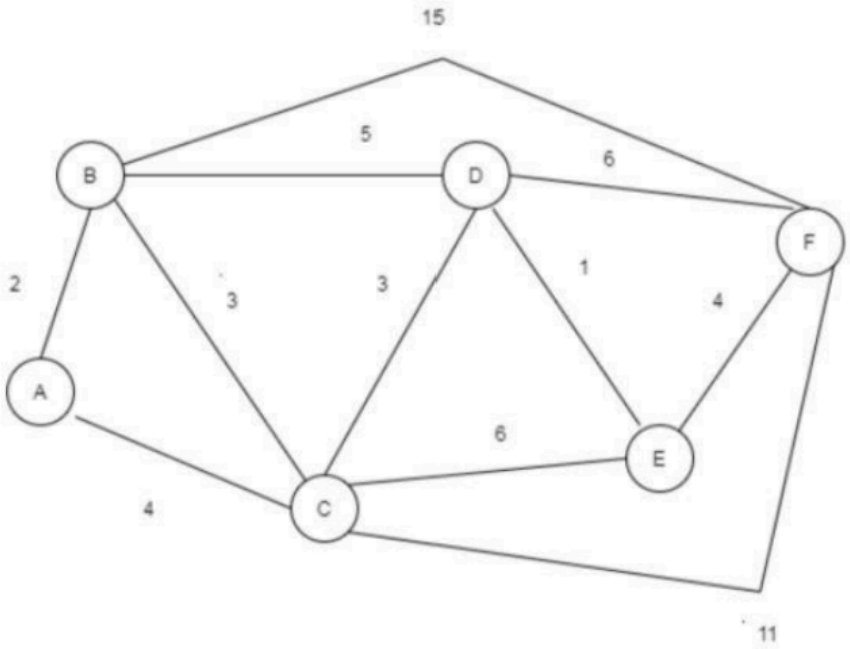
	Course Name:	Artificial Intelligence	Course Code:	AI2002
	Program:	BS Computer Science	Semester:	Sp 2025
	Section	BSC-6H	Total Marks:	50
	Due Date:		Weightage:	3.33
	Exam Type:	Assignment 2	Page(s):	2

Instructions:

- Do the work by yourself, this is an individual assignment.
- Plagiarism cases will be dealt with strictly.
- Late submissions will not be acceptable.

Question #1: [20] [Handwritten Solution (upload photos in GCR)]

Suppose you have been given a map of 6 cities connected with each other via different paths. Your job is to visit every city just once covering the minimum distance possible. Solve this problem using **Genetic Algorithm**. You can start at any point and end at any point. Just make sure that all the cities have been covered.



1. Encode the problem and create an initial population of 3 different chromosomes and choose any

one parent from your above solution and identity the following:

- Gene
- Chromosome

2. Think of an appropriate fitness function to this problem and give proper justification.

3. Use the fitness function to calculate the fitness level of all the chromosomes in your population. Select the fittest 2 chromosomes based on the fitness function.

4. Perform crossover that you have been taught in the class on the selected parents. Now based on the offspring, for this problem do you think that is the best way to perform crossover? If not, explain why.

5. Perform mutation that you have been taught in the class on the produced offspring. Now based on the mutated offspring, for this problem do you think that is the best way to perform mutation? If not, explain why.

Question #2: [30]

In this assignment, you will apply **Linear Regression** to analyze the relationship between an employee's **years of experience** and their **salary**. You will train a regression model, evaluate its performance, and use it to make predictions. **Submit your Python script `(rollNo.ipynb file)`**

A dataset (Salary_dataset.csv) is provided, containing the following columns:

- YearsExperience (X): The number of years of work experience.
- Salary (Y): The corresponding salary of the employee.

Tasks

1. **Load the Dataset:**
    - Read the dataset into a Pandas DataFrame.
    - Display the first five rows to understand its structure.
2. **Data Visualization:**
    - Plot a scatter plot of YearsExperience vs. Salary.
    - What kind of relationship do you observe? (linear, non-linear, etc.)
3. **Train a Linear Regression Model:**
    - Split the data into **training (80%)** and **testing (20%)** sets.
    - Train a **Simple Linear Regression** model using **Scikit-learn**.
4. **Model Evaluation:**
    - Compute the **Mean Squared Error (MSE)** of the model on the test data.
    - Compute the **R^2 score** to evaluate model performance.
    - Interpret what the **R^2 score** tells about the model's accuracy.