UPES
End Semester Examination, May 2024

Course: Machine Learning and Artificial Intelligence in Sustainability   Semester: IV
Program: B. Tech Sustainability Engineering   Time: 03 hrs.
Course Code: SUEN 2005   Max. Marks: 100

Instructions: Attempt all the questions

SECTION A
(5Qx4M=20Marks)

S. No.		Marks	CO
Q 1	Construct a **bar graph** of the following data: 	Marks	<20	<40	<60	<80	<100	 No of Students	10	40	80	100	110		4	CO1
Q 2	The following are the monthly evaporation data (Jan.-Dec.) at a Dam in a certain year in cm: **16, 14, 17, 25, 28, 21, 16, 16, 16, 21, 16, 16** Calculate the **variance** for the data and interpret its physical significance.	4	CO1
Q 3	Describe the design of an intelligent agent.	4	CO2
Q 4	Define the following terms: a) Intelligent Agents b) Stochastic Process c) Effectors d) Blackbox Model	4	CO3
Q 5	Differentiate between the following: a) Supervised and reinforced learning b) Deterministic and stochastic model	4	CO4

SECTION B
(4Qx10M= 40 Marks)

Q 6	A sample of **10** is drawn randomly from a certain population. The sum of the squared deviations from the mean of the given sample is **50**. Test the hypothesis that the variance of the population is **5** at **5** percent level of significance. **OR** There are **100** students in a university college and the whole university, inclusive of this college, the number of students is **2000**. In a random sample study, **20** were found smokers in the college and the proportion of smokers in the university was **0.05**. Is there a significant difference between the proportion of smokers in college and university? Test at **5** percent confidence level.	**10**	**CO1**
Q 7	Write a Python program to perform clustering (dendrograms) to combine the states of India having the water quality parameters data: **Nitrate, fluoride, rain, soil type, and groundwater levels**. Explain each line of code.	**10**	**CO4**
Q 8	Analyze the applications of AI & ML in sustainability. Elucidate with examples.	**10**	**CO3**
Q 9	Why **Monte Carlo Simulations** are important in the case of a stochastic process? Analyze the dice ball game (as in Casino) to predict the chances of the player winning the game.	**10**	**CO2**
SECTION-C **(2Qx20M=40 Marks)**		
Q 10	Describe the popularity of ANN models in machine learning. Present the general architecture of ANN models. Design an ANN model for water potability prediction using the following parameters: **pH, Hardness, Chloramines, Sulphate, Conductivity, Organic carbon, Trihalomethanes, and Turbidity.** **OR** The stage-discharge data of a river are given below. Establish the stage-discharge relationship to predict the discharge for a given stage. Assume the value of stage for zero discharge as 35.00 m. What is the correlation coefficient of the relationship established above? Estimate the discharge corresponding to stage values of **40** m and **48** m respectively.	**20**	**CO4**

Stage (m)	Discharge (m³/s)	Stage (m)	Discharge (m³/s)
35.91	89	39.07	469
36.90	230	41.00	798
37.92	360	43.53	2800
44.40	3800	48.02	5900
45.40	4560	49.05	6800
46.43	5305	49.55	6900
		49.68	6950

The relationship between the discharge (Q) and gauge reading (G) is non-linear and given as:

$$Q = C_r.\,(G - a)^b$$

a = constant which represents the gauge reading corresponding to zero discharge, C_r and b are rating curve constants.

| Q 11 | For the data given below, plot the series and construct an **ARMA** model to predict the future values. | 20 | CO3 |

Month	Passengers	Month	Passengers
Jan-49	112	Jan-50	115
Feb-49	118	Feb-50	126
Mar-49	132	Mar-50	141
Apr-49	129	Apr-50	135
May-49	121	May-50	125
Jun-49	135	Jun-50	149
Jul-49	148	Jul-50	170
Aug-49	148	Aug-50	170
Sep-49	136	Sep-50	158
Oct-49	119	Oct-50	133
Nov-49	104	Nov-50	114
Dec-49	118	Dec-50	140