Name: Enrolment No:	⚡UPES UNIVERSITY OF TOMORROW

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, May 2022

Course: Machine Learning   **Semester: VI**
Program: B.Tech(CSE -Spl. BAO./ Big Data/ Blockchain Technology/ DevOps)   **Time   : 03 hrs.**
Course Code: CSAI 2001P   **Max. Marks: 100**

Instructions: All the Best☺

<table>
<tr><td colspan="4" align="center">SECTION A
(5Qx4M=20Marks)</td></tr>
<tr><td>S. No.</td><td></td><td>Marks</td><td></td></tr>
<tr>
<td>Q 1</td>
<td>Fill in the Blanks using following words:

average linking, Cosine, 1, Elbow, Cluster Analysis,
Support vector machines, Probably Approximately Correct, Multivariate,
Direct regression method, Ordinary Least Square,
one variable in terms of another, coefficient of determination, Machine Learning, Dichotomous, Association rule learning, Minimum,
Supervised Learning, Degree of freedom,
Coefficient of determination, Hierarchical clustering,
Adjusted R-squared, Ensemble, 1NN

a. _____learning is a learning framework that has been introduced to analyze learning algorithms and their statistical efficiency.
b. The main purpose(s) of (LR) is/are (choose all that apply) is to explain ____.
c. In MLR, the square of the multiple correlation coefficient or $R2$ is called the _____.
d. _____ is a learning from labeled data using classification and regression models.
e. _____is considered as the number of observations in regression analysis.
f. The percent of total variation of the dependent variable Y explained by the set of independent variables X is measured by _____.
g. ………………is sub branch of Artificial Intelligence.</td>
<td>4</td>
<td>CO1</td>
</tr>
</table>

	h. ………………..methods extract the rules which efficiently represent the observed relationships between variables in data.		
Q2	Define SSE and SSR	**4**	**CO2**
Q3	Compare dependent and independent variables?	**4**	**CO3**
Q4	Write List of the different type of clusters.	**4**	**CO4**
Q5	Illustrate and Define entropy.	**4**	**CO5**
SECTION B **(4Qx10M= 40 Marks)**			
Q6	Compare supervised learning to unsupervised learning.	**10**	**CO1**
Q7	Describe the procedure of testing of significance using t test. What is the significance of t-test?	**10**	**CO3**
Q8	Explain MLP with an example. Or Illustrate and discuss the procedure of constructing decision tree with an example.	**10**	**CO5**
Q9	Discuss in brief different types of regression models.	**10**	**CO4**
SECTION-C **(2Qx20M=40 Marks)**			
Q10	Analyze and define over-fitting. Give an example. Solve and discuss 4 real-world application for machine Learning. Calculate confusion matrix. Illustrate accuracy with a confusion matrix example.	**6+7+7=20**	**CO2,5**
Q11	Analyze DBSCAN algorithm. Comment on its time complexity. Explain k-means clustering technique. Discuss bisecting k-means technique. Or Discuss SVM as a classifier. Compare SVM with KNN and elaborate about them using diagrams.	**20**	**CO4**