

Name:	
Enrolment No:	

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

Course: Cognitive Analytics
Program: B.Tech (CSE)-AIML
Course Code: CSBA3009

Semester: 6
Time: 03 hrs.
Max. Marks: 100

Instructions: Attempt all questions.

SECTION A
(5Qx4M=20Marks)

S. No.		Marks	CO
Q 1	What is a cognitive system? Describe the characteristics of a cognitive system.	4	CO1
Q 2	Differentiate between simple random sampling and stratified sampling with suitable examples.	4	CO1
Q 3	Write R programming code to implement multiple linear regression from the scratch.	4	CO2
Q 4	Precisely describe Bias and Variance Tradeoff. Define Overfitting and Underfitting in terms of bias and variance.	4	CO2
Q 5	How cloud computing is beneficial in cognitive analytics services?	4	CO3

SECTION B
(4Qx10M= 40 Marks)

Q 6	<p>The table below shows year-wise sales data of different countries for various markets for years 2012 to 2015.</p> <table border="1" style="width: 100%; text-align: center;"> <thead> <tr> <th>Year</th> <th>2012</th> <th>2013</th> <th>2014</th> <th>2015</th> </tr> </thead> <tbody> <tr> <td>India</td> <td>127187.27</td> <td>144480.7</td> <td>229068.79</td> <td>283036.44</td> </tr> <tr> <td>Sri Lanka</td> <td>492756.6</td> <td>486629.3</td> <td>627634.98</td> <td>757108.13</td> </tr> <tr> <td>Nepal</td> <td>385098.15</td> <td>464733.29</td> <td>608140.77</td> <td>706632.93</td> </tr> <tr> <td>Pakistan</td> <td>713658.22</td> <td>863983.97</td> <td>1092231.65</td> <td>1372784.4</td> </tr> <tr> <td>China</td> <td>540750.63</td> <td>717611.4</td> <td>848670.24</td> <td>1180303.95</td> </tr> </tbody> </table> <p>Write python script for the following visualization operations over the above data:</p> <ul style="list-style-type: none"> (i) To create subplots of 4 columns (ii) To show line graph for China with color red and title “Sales in China”. (iii) To show bar for Sri Lanka with title “Sales in Sri Lanka”. (iv) To show scatter plot for India with title “Sales in India”. 	Year	2012	2013	2014	2015	India	127187.27	144480.7	229068.79	283036.44	Sri Lanka	492756.6	486629.3	627634.98	757108.13	Nepal	385098.15	464733.29	608140.77	706632.93	Pakistan	713658.22	863983.97	1092231.65	1372784.4	China	540750.63	717611.4	848670.24	1180303.95	10	CO1
Year	2012	2013	2014	2015																													
India	127187.27	144480.7	229068.79	283036.44																													
Sri Lanka	492756.6	486629.3	627634.98	757108.13																													
Nepal	385098.15	464733.29	608140.77	706632.93																													
Pakistan	713658.22	863983.97	1092231.65	1372784.4																													
China	540750.63	717611.4	848670.24	1180303.95																													

	(v) To combine data for Pakistan and Nepal and set title "Sales in Pakistan and Nepal". Set color of Pakistan bar to "royalblue" and Nepal bar to sea green and set Legend for them. Set width of Pakistan bar and Nepal bar to 0.5.																																																									
Q 7	Define predictive modelling. Discuss the types of predictive models and the challenges involved in predictive modelling.	3+4+3=10	CO1																																																							
Q 8	How density based clustering is different from hierarchical clustering? Write down the algorithm for DBSCAN clustering method.	5+5=10	CO2																																																							
Q 9	How cloud computing is beneficial in cognitive analytics? What specific challenges it will be able to meet and how?	10	CO3																																																							
SECTION-C (2Qx20M=40 Marks)																																																										
Q 10	<p>Describe the principle of Naïve Bayesian Classifier with its advantages and disadvantages.</p> <p>Apply Naïve Bayesian Classifier over the data given below and find the Evade class for the following instance:</p> <p>$X = (Refund=No, Marital\ Status=Married, Income=120K)$</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Tid</th> <th>Refund</th> <th>Marital Status</th> <th>Income</th> <th>Evade</th> </tr> </thead> <tbody> <tr><td>1</td><td>Yes</td><td>Single</td><td>125K</td><td>No</td></tr> <tr><td>2</td><td>No</td><td>Married</td><td>100K</td><td>No</td></tr> <tr><td>3</td><td>No</td><td>Single</td><td>70K</td><td>No</td></tr> <tr><td>4</td><td>Yes</td><td>Married</td><td>120K</td><td>No</td></tr> <tr><td>5</td><td>No</td><td>Divorced</td><td>95K</td><td>Yes</td></tr> <tr><td>6</td><td>No</td><td>Married</td><td>60K</td><td>No</td></tr> <tr><td>7</td><td>Yes</td><td>Divorced</td><td>220K</td><td>No</td></tr> <tr><td>8</td><td>No</td><td>Single</td><td>85K</td><td>Yes</td></tr> <tr><td>9</td><td>No</td><td>Married</td><td>75K</td><td>No</td></tr> <tr><td>10</td><td>No</td><td>Single</td><td>90K</td><td>Yes</td></tr> </tbody> </table>	Tid	Refund	Marital Status	Income	Evade	1	Yes	Single	125K	No	2	No	Married	100K	No	3	No	Single	70K	No	4	Yes	Married	120K	No	5	No	Divorced	95K	Yes	6	No	Married	60K	No	7	Yes	Divorced	220K	No	8	No	Single	85K	Yes	9	No	Married	75K	No	10	No	Single	90K	Yes	6+14=20	CO2
Tid	Refund	Marital Status	Income	Evade																																																						
1	Yes	Single	125K	No																																																						
2	No	Married	100K	No																																																						
3	No	Single	70K	No																																																						
4	Yes	Married	120K	No																																																						
5	No	Divorced	95K	Yes																																																						
6	No	Married	60K	No																																																						
7	Yes	Divorced	220K	No																																																						
8	No	Single	85K	Yes																																																						
9	No	Married	75K	No																																																						
10	No	Single	90K	Yes																																																						
Q 11	<p>What are the benefits of machine learning with IBM's Watson over traditional methods of machine learning with python or R?</p> <p>Define core capabilities of IBM's Watson.</p> <p>Describe the working of DeepQA with its system architecture.</p>	5+5+10=20	CO3																																																							