



GSSS INSTITUTE OF ENGINEERING & TECHNOLOGY FOR WOMEN

KRS Road, Metagalli, Mysuru., 570016

(Affiliated to VTU, Belagavi, Approved by AICTE, New Delhi & Govt of Karnataka)

(Accredited with Grade 'A' by NAAC)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

(Accredited by NBA, New Delhi, Validity: 01.07.2017 – 30.06.2020 & 01.07.2020 – 30.06.2023)

SECOND ONLINE INTERNAL ASSESSMENT TEST, EVEN SEM, 2020-2021

SEMESTER: 8th 'A'

SUBJECT NAME: **Big data Analytics [17CS82]**

FACULTY NAME: RAJASHEKAR M.B

TIME: **10:30 a.m. - 11:40 a.m**

DATE: 29/06/2021

TOTAL MARKS: **30 Marks**

Instructions to the students: Answer any one question from each part.

Course Outcome:

CO3: Recognize the role of Business Intelligence, Data warehousing and Visualization in decision making

CO4: Infer the importance of core data mining techniques for data analytics

Q. No		Marks	CO	Blooms Taxonomy																																																																																					
PART A																																																																																									
1(a)	Explain CRISP-DM Cycle with a neat diagram.	5M	CO3	Understand																																																																																					
(b)	Explain the star schema design of Data Warehousing with an example	5M	CO3	Understand																																																																																					
OR																																																																																									
2(a)	Differentiate the following i) Data Mining and Data Warehousing ii) Supervised learning and Unsupervised Learning	5M	CO3	Analyze																																																																																					
(b)	Describe Business Intelligence Tools with example.	5M	CO3	Understand																																																																																					
PART B																																																																																									
3(a)	Define Splitting Variable. Describe three criteria for choosing a Splitting variable?	5M	CO4	Understand																																																																																					
(b)	List some of the advantages and Disadvantages of Regression Model	5M	CO4	Remember																																																																																					
OR																																																																																									
4	Construct the decision tree for the following data set <table border="1"> <thead> <tr> <th>Outlook</th><th>Temp.</th><th>Humidity</th><th>Wind</th><th>Play</th></tr> </thead> <tbody> <tr><td>Sunny</td><td>Hot</td><td>High</td><td>False</td><td>No</td></tr> <tr><td>Sunny</td><td>Hot</td><td>High</td><td>True</td><td>No</td></tr> <tr><td>Overcast</td><td>Hot</td><td>High</td><td>False</td><td>Yes</td></tr> <tr><td>Rain</td><td>Mild</td><td>High</td><td>False</td><td>Yes</td></tr> <tr><td>Rain</td><td>Cool</td><td>Normal</td><td>False</td><td>Yes</td></tr> <tr><td>Rain</td><td>Cool</td><td>Normal</td><td>True</td><td>No</td></tr> <tr><td>Overcast</td><td>Cool</td><td>Normal</td><td>True</td><td>Yes</td></tr> <tr><td>Sunny</td><td>Mild</td><td>High</td><td>False</td><td>No</td></tr> <tr><td>Sunny</td><td>Cool</td><td>Normal</td><td>False</td><td>Yes</td></tr> <tr><td>Rain</td><td>Mild</td><td>Normal</td><td>False</td><td>Yes</td></tr> <tr><td>Sunny</td><td>Mild</td><td>Normal</td><td>True</td><td>Yes</td></tr> <tr><td>Overcast</td><td>Mild</td><td>High</td><td>True</td><td>Yes</td></tr> <tr><td>Overcast</td><td>Hot</td><td>Normal</td><td>False</td><td>Yes</td></tr> <tr><td>Rain</td><td>Mild</td><td>High</td><td>True</td><td>No</td></tr> </tbody> </table> <p>Then solve the following problem using the model</p> <table border="1"> <thead> <tr> <th>Outlook</th><th>Temp</th><th>Humidity</th><th>Windy</th><th>Play</th></tr> </thead> <tbody> <tr> <td>Sunny</td><td>Hot</td><td>Normal</td><td>True</td><td>???</td></tr> </tbody> </table>	Outlook	Temp.	Humidity	Wind	Play	Sunny	Hot	High	False	No	Sunny	Hot	High	True	No	Overcast	Hot	High	False	Yes	Rain	Mild	High	False	Yes	Rain	Cool	Normal	False	Yes	Rain	Cool	Normal	True	No	Overcast	Cool	Normal	True	Yes	Sunny	Mild	High	False	No	Sunny	Cool	Normal	False	Yes	Rain	Mild	Normal	False	Yes	Sunny	Mild	Normal	True	Yes	Overcast	Mild	High	True	Yes	Overcast	Hot	Normal	False	Yes	Rain	Mild	High	True	No	Outlook	Temp	Humidity	Windy	Play	Sunny	Hot	Normal	True	???	10M	CO4	Apply
Outlook	Temp.	Humidity	Wind	Play																																																																																					
Sunny	Hot	High	False	No																																																																																					
Sunny	Hot	High	True	No																																																																																					
Overcast	Hot	High	False	Yes																																																																																					
Rain	Mild	High	False	Yes																																																																																					
Rain	Cool	Normal	False	Yes																																																																																					
Rain	Cool	Normal	True	No																																																																																					
Overcast	Cool	Normal	True	Yes																																																																																					
Sunny	Mild	High	False	No																																																																																					
Sunny	Cool	Normal	False	Yes																																																																																					
Rain	Mild	Normal	False	Yes																																																																																					
Sunny	Mild	Normal	True	Yes																																																																																					
Overcast	Mild	High	True	Yes																																																																																					
Overcast	Hot	Normal	False	Yes																																																																																					
Rain	Mild	High	True	No																																																																																					
Outlook	Temp	Humidity	Windy	Play																																																																																					
Sunny	Hot	Normal	True	???																																																																																					

PART C				
5(a)	Describe advantages & disadvantages of K-Means algorithm	5M	CO4	Understand
(b)	Explain Elbow method for determining number of clusters in a data set with neat diagram.	5M	CO4	Understand
OR				
6(a)	Briefly explain the data mining mistakes	5M	CO3	Analyze
(b)	With diagram explain the model for Multilayer ANN	5M	CO4	Analyze

RAJASHEKAR M.B
Faculty signature

Review committee

HOD/CSE

Dr. Raviraj P
Professor, Dept of CSE

Chaya P
Asst. professor, Dept of ISE