U.S.N | | | | | | | | | | |

P.E.S. College of Engineering, Mandya - 571 401

(An Autonomous Institution affiliated to VTU, Belagavi)

Seventh Semester, B.E. - Computer Science and Engineering

Semester End Examination; February - 2022

Data Analytics

Time: 3 hrs *Max. Marks: 100*

Course Outcomes

The Students will be able to:

Co1: Understanding the Data Analytics Taxonomy and Descriptive Statistics:
CO2: Apply data processing methods.
CO3: Analyze clustering and Frequent Pattern Mining algorithms.
Co4: Analyze regression and classification algorithms.
CO5: Design and Implement real time applications in data analytics.

Note: ***I) PART - A*** *is compulsory.* ***Two*** *marks for each question.*
II) PART - B*: Answer any* ***Two*** *sub questions (from a, b, c) for Maximum of* ***18 marks*** *from each unit.*

Q. No.		Questions	Marks	BLs	COs	POs
		I : PART - A	**10**			
I	a.	What is small data?	2	L2	CO1	PO1,2,3
	b.	List the approaches used to deal with a data set which contains missing values.	2	L2	CO2	PO1,2,3
	c.	What is Clustering?	2	L2	CO3	PO1,2,3
	d.	List any four domains where the regression can be used.	2	L2	CO4	PO1,2,3
	e.	What are rule-set induction algorithms?	2	L2	CO5	PO1,2,3,4
		II : PART - B	**90**			
		UNIT - I	**18**			
1	a.	Explain the taxonomy of data analytics with an example.	9	L2	CO1	PO1,2,3
	b.	Explain the steps involved in the KDD process.	9	L2	CO1	PO1,2,3
	c.	List the various scale types used in data analytics. Explain the same with suitable examples.	9	L2	CO1	PO1,2,3
		UNIT - II	**18**			
2	a.	With an example, illustrate the process to be followed while working with noisy data and outliers.	9	L3	CO2	PO1,2,3,4
	b.	Explain the conversion from nominal scale to relative scale with suitable example.	9	L2	CO2	PO1,2,3,4
	c.	Explain the process of Principal Component Analysis (PCA) with suitable example.	9	L2	CO2	PO1,2,3,4

Contd… 2

	UNIT - III	**18**			
3 a.	With an algorithm, explain K-Means clustering technique.	9	L2	CO3	PO2
b.	Explain Agglomerative Hierarchical clustering technique with suitable example.	9	L2	CO3	PO2
c.	Explain the Apriori algorithm for frequent item sets generation.	9	L2	CO3	PO2
	UNIT - IV	**18**			
4 a.	Explain the "Random Sub-Sampling" and "K-Fold cross validation" of model validation with suitable example.	9	L2	CO4	PO1,2,3
b.	With a suitable diagram, explain the confusion matrix and predictive performance measures used in binary classification.	9	L2	CO4	PO1,2,3
c.	With an algorithm, explain K-Nearest Neighbor algorithm.	9	L2	CO4	PO1,2,3
	UNIT - V	**18**			
5 a.	Explain the five phases of text mining.	9	L2	CO5	PO1,2,3,4
b.	Explain the hunt decision tree induction algorithm with example.	9	L2	CO5	PO1,2,3,4
c.	Explain Support Vector Machines with advantages and disadvantages.	9	L2	CO5	PO1,2,3,4

* * * *