B.Tech
PCS4D001

4ᵗʰ Semester Regular/Back Examination 2018-19
DATA ANALYTICS
BRANCH : CSE
Max Marks : 100
Time : 3 Hours
Q.CODE : F954
Answer Question No.1 (Part-1) which is compulsory, any eight from Part-II and any two from Part-III.
The figures in the right hand margin indicate marks.

Part- I

Q1	**Only Short Answer Type Questions (Answer All-10)**	**(2 x 10)**
a)	Mention some common problems that data analysts encounter during analysis.	
b)	How will you create a classification to identify key customer trends in unstructured data?	
c)	What is the criteria to say whether a developed data model is good or not?	
d)	Define multilayer classifier.	
e)	Define the Bayes classifier.	
f)	What's the difference between a generative and discriminative model?	
g)	In which problem the logistic regression is used?	
h)	What is BigData Analysis?	
i)	Compare between probability and likelihood	
j)	What is data cleansing?	

Part- II

Q2	**Only Focused-Short Answer Type Questions- (Answer Any Eight out of Twelve)**	**(6 x 8)**
a)	Explain the typical data analysis process.	
b)	Discuss the issues and challenges of KNN model.	
c)	Discuss details on cross validation method	
d)	Why subset selection method is essential. Discuss the best selection method with its demerits.	
e)	Discuss the main features of data analytics method in details.	
f)	Compare between bootstraps and cross validation with proper examples.	
g)	How ANN learning in data analysis provide better result.	
h)	What cross-validation technique would you use on a time series dataset? Justify your answer with example.	
i)	How would you handle an imbalanced dataset?	
j)	What is the role of kernel machine?	
k)	What is use of forward propagation in neural network?	
l)	How would you evaluate a logistic regression model?	

Part-III

Only Long Answer Type Questions (Answer Any Two out of Four)

Q3	What is SVM? Mention the role of SVM. Explain the linear classifier related to SVM	**(16)**
Q4	Differentiate between supervised and unsupervised learning in details with proper examples.	**(16)**
Q5	Discussed the different components of Artificial Neural Networks with examples.	**(16)**
Q6 a)	How best use of least-squared error hypotheses?	**(8)**
b)	Discuss the Decision tree representation in data analysis.	**(8)**