CBCS SCHEME

USN | | | | | | | | | | |

21AI743

Seventh Semester B.E./B.Tech. Degree Examination, Dec.2024/Jan.2025
Predictive Analytics

Time: 3 hrs. Max. Marks: 100

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Module-1

1 a. Define predictive modeling. Describe its importance in Business Analytics. **(06 Marks)**

b. Explain how propensity models are used in predictive analytics with an example. **(08 Marks)**

c. Discuss the types of Business Analytics in brief. **(06 Marks)**

OR

2 a. Identify and explain three key challenges in developing accurate predictive models. **(06 Marks)**

b. Describe overfitting in predictive modeling and suggest strategies to mitigate it. **(08 Marks)**

c. Discuss the role of clustering in customer segmentation. **(06 Marks)**

Module-2

3 a. Explain the differences between statistical and machine learning models. **(06 Marks)**

b. Discuss the significance of hyper parameter tuning in machine learning. **(06 Marks)**

c. Compare decision trees and random forests in predictive modeling. **(08 Marks)**

OR

4 a. Define overfitting and underfitting. Explain how cross-validation helps addressing their issues. **(08 Marks)**

b. Explain the benefits of ensemble methods in predictive modeling. **(06 Marks)**

c. Describe how Support Vector Machine (SVM) can be used for both regression and classification. **(06 Marks)**

Module-3

5 a. Explain feature engineering and its significance in predictive modeling. **(06 Marks)**

b. Compare and contrast data transformation and data normalization. **(06 Marks)**

c. Describe missing data handling and list with brief, common imputation techniques. **(08 Marks)**

OR

6 a. Explain the importance of data scaling in models like SVM and k-NN. **(08 Marks)**

b. Describe predictor binning in data preprocessing. **(04 Marks)**

c. Write an "R" program to apply logarithmic transformation to skewed data.

Store ID	Wait Time (minutes)
A	5
B	6
C	5
D	4
E	7
F	12
G	6
H	11

(08 Marks)

Module-4

7 a. Explain the differences between linear and nonlinear regression models with examples. **(08 Marks)**

b. Discuss the assumptions underlying linear regression and the impact of assumption violation. **(06 Marks)**

c. Write an 'r' program to create a linear regression model and summarize its output for the given data. **(06 Marks)**

Data for Q.7(c) and Q.9(c)

Row	Spend	Web Traffic	Discount	Sales
1	1629	2383	21.3	1877.5
2	1332	1842	14.5	3425.4
3	1894	4231	23.7	3076.8
4	697	1221	10.2	1533.9
5	1268	2108	27.4	1892.1
6	1986	4634	15.9	5496.6
7	612	1443	25.3	1227.8
8	1749	3782	12.7	4512.5

OR

8 a. Define multicollinearity and its impact on regression models. **(08 Marks)**

b. Discuss the bias-variance trade off in regression model. **(06 Marks)**

c. Explain r-squared and its uses in evaluation of model performance. **(06 Marks)**

Module-5

9 a. Compare logistic regression with discriminant analysis for classification tasks. **(08 Marks)**

b. Describe confusion matrix and its role in model evaluation. **(06 Marks)**

c. Write a "R" program to perform logistic regression on a dataset and evaluate its performance. **(06 Marks)**

OR

10 a. Explain the difference between precision and recall in classification. **(04 Marks)**

b. Discuss the role of feature selection and how it improves model accuracy. **(06 Marks)**

c. Explain how Receiver Operating Characteristic (ROC) curves evaluate classification model performance with an example. **(10 Marks)**

* * * * *