

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BSc DEGREE EXAMINATION DECEMBER 2022
(Fifth Semester)

Branch – COMPUTER SCIENCE

DISCIPLINE SPECIFIC ELECTIVE: I – PREDICTIVE ANALYTICS

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer ALL questions

ALL questions carry EQUAL marks (10 x 1 = 10)

1. Supervised learning is also called _____
(i) Predictive Modeling (ii) Descriptive Modeling
(iii) Predictive Data (iv) Text Modeling
2. The output of business intelligence analyses are reports or dashboards that summarize interesting characteristics of the data, is called as _____
(i) Key Performance Interruption (ii) Key Prediction Indicators
(iii) Key Performance Indicators (iv) Key Predictive Indicators
3. The _____ value represent the middle of the distribution or a typical value
(i) Median (ii) Mean
(iii) Standard Deviation (iv) Distribution
4. _____ is a graphical representation of the quartile statistics of a variable and a method insight into the characteristics of numeric data.
(i) Spike (ii) Correlation
(iii) Interaction (iv) Box Plot
5. Changing values of missing data to a value that represents an expected value in the variable if it were actually known is called as _____
(i) Missing value imputation (ii) List Wise Deletion
(iii) Column Wise Deletion (iv) Row Deletion
6. _____ finds linear projections of numeric data that maximize the spread or variables of the projections
(i) Principal Component Analysis (ii) Principal Complex Analysis
(iii) Linear Analysis (iv) Regression Analysis
7. _____ is defined as the number of times a rule occurs in the data divided by the number of transactions in the data
(i) Confidence (ii) Consequent
(iii) Support (iv) Antecedent
8. Association rules operate only on _____ data.
(i) Target (ii) Pattern
(iii) Continuous (iv) Categorical
9. _____ models are interpreted by examining the conditional probabilities generated in the training data
(i) KNN (ii) Naïve Bayes
(iii) Regression (iv) Logistic

Cont...

- 10 The simplest method of variable selection is called _____ operates in variable selection occurs in decision trees
- (i) Backward Selection (ii) Networks selection
(iii) Forward Selection (iv) Probability Selection

SECTION - B (25 Marks)

Answer ALL questions

ALL questions carry EQUAL Marks (5 x 5 = 25)

- 11 a Business Intelligence “look back the past” and Predictive Analytics “predicts the future” – Justify in your own words.
OR
b Explain the three-legged stool in the predictive modeling.
- 12 a Show the three distributions in kurtosis with a neat sketch.
OR
b Sketch the scatter plot and its matrix with an example.
- 13 a Identify the identification of the missing values and fixing the missing data in variable cleaning process.
OR
b Describe the variable selection prior to the modeling.
- 14 a List out the steps to build a classification rule from Association rules.
OR
b Narrate the importance of KNN algorithm.
- 15 a Show the variable selection in Linear Regression.
OR
b Outline the metrics in assessing Regression Models.

SECTION -C (40 Marks)

Answer ALL questions

ALL questions carry EQUAL Marks (5 x 8 = 40)

- 16 a Elaborate the supervised and unsupervised learning with a relevant example.
OR
b Compare and contrast Predictive Analytics and Business Intelligence.
- 17 a Elucidate the importance of box plot graphical representation.
OR
b “Scatterplots are the most commonly used Two-Dimensional visualization method”- comment the statement and discuss the importance of four scatterplots - Anscombe’s Quartet.
- 18 a Summarise the Missing value codes, its types, list wise and column deletion methods in handling missing values.
OR
b Discuss the significance of variable transformations in feature creation.
- 19 a Enumerate the association rule with a common application example.
OR
b Highlight the significant features of Logistic Regression with an example.
- 20 a “Linear regression varies from logistic regression” – Justify.
OR
b Discuss the assessment of predictive model with a suitable example.