

Roll. No.

Question Booklet Number

O.M.R. Serial No.

--	--	--	--	--	--	--	--

BCA (SEM.-VI) (NEP) (SUPPLE.) EXAMINATION, 2024-25

COMPUTER APPLICATION

(Data Science & Machine Learning)

Paper Code							
Z	0	1	0	1	2	0	T

[BCA-6004]

Question Booklet
Series

A

Time : 1 : 30 Hours

Max. Marks : 75

Instructions to the Examinee :

1. Do not open the booklet unless you are asked to do so.
2. The booklet contains 100 questions. Examinee is required to answer 75 questions in the OMR Answer-Sheet provided and not in the question booklet. All questions carry equal marks.
3. Examine the Booklet and the OMR Answer-Sheet very carefully before you proceed. Faulty question booklet due to missing or duplicate pages/questions or having any other discrepancy should be got immediately replaced.
4. Four alternative answers are mentioned for each question as - A, B, C & D in the booklet. The candidate has to choose the correct / answer and mark the same in the OMR Answer-Sheet as per the direction :

(Remaining instructions on last page)

परीक्षार्थियों के लिए निर्देश :

1. प्रश्न-पुस्तिका को तब तक न खोलें जब तक आपसे कहा न जाए।
2. प्रश्न-पुस्तिका में 100 प्रश्न हैं। परीक्षार्थी को 75 प्रश्नों को केवल दी गई OMR आन्सर-शीट पर ही हल करना है, प्रश्न-पुस्तिका पर नहीं। सभी प्रश्नों के अंक समान हैं।
3. प्रश्नों के उत्तर अंकित करने से पूर्व प्रश्न-पुस्तिका तथा OMR आन्सर-शीट को सावधानीपूर्वक देख लें। दोषपूर्ण प्रश्न-पुस्तिका जिसमें कुछ भाग छपने से छूट गए हों या प्रश्न एक से अधिक बार छप गए हों या उसमें किसी अन्य प्रकार की कमी हो, उसे तुरन्त बदल लें।
4. प्रश्न-पुस्तिका में प्रत्येक प्रश्न के चार सम्भावित उत्तर- A, B, C एवं D हैं। परीक्षार्थी को उन चारों विकल्पों में से सही उत्तर छानना है। उत्तर को OMR उत्तर-पत्रक में सम्बन्धित प्रश्न संख्या में निम्न प्रकार भरना है :

(शेष निर्देश अन्तिम पृष्ठ पर)

1. Data Science is primarily concerned with:
 - (A) Data storage
 - (B) Extracting knowledge and insights from data
 - (C) Data deletion
 - (D) Data encryption
2. Which of the following is not a role in Data Science?
 - (A) Data Engineer
 - (B) Data Scientist
 - (C) Data Doctor
 - (D) Data Analyst
3. The first stage in a Data Science project is:
 - (A) Data Modeling
 - (B) Data Visualization
 - (C) Problem Definition
 - (D) Model Deployment
4. Which of these is an application of Data Science?
 - (A) Healthcare diagnosis
 - (B) Fraud detection
 - (C) Recommendation systems
 - (D) All of the above
5. The term "Big Data" is often characterized by:
 - (A) Volume, Variety, Velocity
 - (B) Value, Velocity, Vividness
 - (C) Volume, Variance, Visibility
 - (D) Variation, Volume, Visuals
6. Data Science overlaps with which of the following fields?
 - (A) Statistics
 - (B) Computer Science
 - (C) Domain Knowledge
 - (D) All of the above
7. A major challenge in Data Science is:
 - (A) Lack of storage
 - (B) Data Security Issues
 - (C) Limited programming languages
 - (D) Absence of statistics
8. A person who develops and maintains data pipelines is called:
 - (A) Data Scientist
 - (B) Data Engineer
 - (C) Data Analyst
 - (D) Data Manager

9. The goal of Data Science is to:
- (A) Collect as much data as possible
 - (B) Clean all types of data
 - (C) Extract actionable insights
 - (D) Visualize only structured data
10. Which of the following is not a stage in a Data Science project?
- (A) Data Collection
 - (B) Model Training
 - (C) Web Development
 - (D) Data Cleaning
11. An example of Data Science in finance is:
- (A) Predicting stock market trends
 - (B) Creating comic books
 - (C) Developing musical instruments
 - (D) Growing crops
12. In Data Science, "Data Security" refers to:
- (a) Storing data in a spreadsheet
 - (B) Protecting data from unauthorized access
 - (C) Deleting unnecessary files
 - (D) Using data only for visualization
13. Which of these best describes a Data Scientist?
- (A) Only a programmer
 - (B) Only a statistician
 - (C) A professional combining statistics, programming, and domain knowledge
 - (D) A database administrator
14. One of the key applications of Data Science in e-commerce is;
- (A) Customer recommendation engines
 - (B) Painting artwork
 - (C) Building highways
 - (D) None of the above
15. Which role is responsible for analyzing data trends and patterns?
- (A) Data Engineer
 - (B) Data Analyst
 - (C) Network Engineer
 - (D) Web Developer
16. Which company is famous for pioneering Data Science in personalized recommendations?
- (A) Netflix
 - (B) Ford
 - (C) Nokia
 - (D) Philips

17. A key challenge in Data Science projects is:
 - (A) Lack of programming languages
 - (B) Handling unstructured data
 - (C) No need for visualization tools
 - (D) Eliminating statistics
18. Data Science contributes to healthcare by:
 - (A) Diagnosing diseases using patient data
 - (B) Manufacturing medicines
 - (C) Building hospitals
 - (D) Supplying medical equipment
19. Which is not an example of Data Science application?
 - (A) Spam filtering in emails
 - (B) Image recognition
 - (C) Social media analytics
 - (D) Manual bookkeeping
20. Data Science has evolved from which main disciplines?
 - (A) Physics and Biology
 - (B) Statistics, Computer Science, and Data Engineering
 - (C) Music and Arts
 - (D) Architecture and Civil Engineering
21. Which of the following is the first step in Data Preprocessing?
 - (A) Data Integration
 - (B) Data Cleaning
 - (C) Data Reduction
 - (D) Data Visualization
22. The process of combining data from multiple sources is called:
 - (A) Data Cleaning
 - (B) Data Integration
 - (C) Data Transformation
 - (D) Data Reduction
23. Handling missing values Is part of:
 - (A) Data Collection
 - (B) Data Cleaning
 - (C) Data Modeling
 - (D) Data Deployment
24. Which of these is not a data collection strategy?
 - (A) Surveys
 - (B) Experiments
 - (C) Random guessing
 - (D) Observations

25. Data Transformation is required to:
- (A) Convert data into a suitable format
 - (B) Increase data size
 - (C) Delete Irrelevant data
 - (D) Reduce redundancy
26. Which of the following methods is used to reduce data size while maintaining integrity?
- (A) Data Cleaning
 - (B) Data Reduction
 - (C) Data Expansion
 - (D) Data Augmentation
27. An example of a data collection source is:
- (A) Database
 - (B) Social Media
 - (C) IoT devices
 - (D) All of the above
28. Outlier detection is part of:
- (A) Data Reduction
 - (B) Data Cleaning
 - (C) Data Collection
 - (D) Data Visualization
29. Which of these is a Data Preprocessing step?
- (A) Data Modeling
 - (B) Data Integration
 - (C) Cloud Deployment
 - (D) Algorithm Selection
30. Normalization in Data Preprocessing refers to:
- (A) Increasing data redundancy
 - (B) Scaling data into a defined range
 - (C) Removing duplicate records
 - (D) Detecting anomalies
31. Which of the following is an example of structured data?
- (A) Tweets
 - (B) Sensor readings in a table
 - (C) Images
 - (D) Videos
32. Data Cleaning includes:
- (A) Removing duplicates
 - (B) Handling missing values
 - (C) Detecting inconsistencies
 - (D) All of the above

33. The process of mapping data into new attributes for better learning is:
- (A) Feature Engineering
 - (B) Data Reduction
 - (C) Data Collection
 - (D) Normalization
34. Which of these is a data reduction technique?
- (A) Dimensionality reduction
 - (B) Sampling
 - (C) Data compression
 - (D) All of the above
35. When raw data is collected without cleaning, it is referred to as:
- (A) Clean Data
 - (B) Dirty Data
 - (C) Normalized Data
 - (D) Processed Data
36. An example of unstructured data is:
- (A) CSV file
 - (B) SQL database
 - (C) Video recordings
 - (D) Excel sheet
37. Data Transformation includes:
- (A) Smoothing
 - (B) Aggregation
 - (C) Generalization
 - (D) All of the above
38. A data preprocessing step that reduces storage and computation costs is:
- (A) Data Expansion
 - (B) Data Reduction
 - (C) Data Transformation
 - (D) Data Visualization
39. Which type of error occurs when irrelevant attributes are included in data?
- (A) Data redundancy
 - (B) Data leakage
 - (C) Data inconsistency
 - (D) Data bias
40. The goal of Data Preprocessing is to:
- (A) Make data ready for analysis
 - (B) Increase data collection
 - (C) Visualize results
 - (D) Deploy a model

41. Exploratory Data Analysis (EDA) primarily focuses on:
- (A) Data Deployment
 - (B) Understanding data patterns and relationships
 - (C) Model evaluation
 - (D) Algorithm selection
42. Which of these is not a descriptive statistic?
- (A) Mean
 - (B) Standard Deviation
 - (C) Regression Coefficient
 - (D) Skewness
43. A measure of how spread out values are around the mean is:
- (A) Mean
 - (B) Standard Deviation
 - (C) Mode
 - (D) Median
44. Skewness is used to measure:
- (A) The central tendency of data
 - (B) The symmetry of the data distribution
 - (C) The spread of data
 - (D) The relationship between two variables
45. If a dataset has high positive skewness, its tail is:
- (A) Toward the left
 - (B) Toward the right
 - (C) Symmetrical
 - (D) None of the above
46. Kurtosis measures:
- (A) The peak and flatness of a distribution
 - (B) The average of a dataset
 - (C) The deviation from the mean
 - (D) The correlation between variables
47. A box plot is used for:
- (A) Showing data distribution and outliers
 - (B) Predicting future values
 - (C) Performing regression analysis
 - (D) Classifying data
48. Which component is not shown in a box plot?
- (A) Median
 - (B) Quartiles
 - (C) Outliers
 - (D) Regression Line

49. Pivot tables are mainly used in:
- (A) Summarizing and analyzing large datasets
 - (B) Model deployment
 - (C) Neural network training
 - (D) Data encryption
50. The correlation coefficient ranges between:
- (A) 0 and 1
 - (B) -1 and +1
 - (C) -2 and +2
 - (D) $-\infty$ and $+\infty$
51. A correlation coefficient close to +1 indicates:
- (A) Strong negative correlation
 - (B) Strong positive correlation
 - (C) No correlation
 - (D) Weak relationship
52. ANOVA is used to:
- (A) Compare means across multiple groups
 - (B) Test data normality
 - (C) Measure correlation
 - (D) Perform clustering
53. In EDA, detecting outliers helps in:
- (A) Improving model accuracy
 - (B) Reducing redundancy
 - (C) Ignoring visualization
 - (D) Deleting all data
54. Which visualization is best for identifying skewness?
- (A) Histogram
 - (B) Scatter Plot
 - (C) Pie Chart
 - (D) Line Graph
55. A scatter plot is used to visualize:
- (A) Distribution of a single variable
 - (B) Relationship between two variables
 - (C) Central tendency
 - (D) Data hierarchy
56. The mean of dataset [5,10,15] is:
- (A) 5
 - (B) 10
 - (C) 15
 - (D) 20

57. The measure of "average squared deviation from the mean" is called:
- (A) Variance
 - (B) Standard Deviation
 - (C) Mean
 - (D) Median
58. Which tool can generate Pivot Tables automatically?
- (A) Python
 - (B) Excel
 - (C) R
 - (D) All of the above
59. A correlation value of 0 implies:
- (A) Strong positive relation
 - (B) Strong negative relation
 - (C) No linear relation
 - (D) Perfect dependency
60. EDA helps in:
- (A) Feature selection
 - (B) Identifying data patterns
 - (C) Outlier detection
 - (D) All of the above
61. Machine Learning is best defined as;
- (A) Writing explicit rules for computers
 - (B) Enabling machines to learn patterns from data
 - (C) Hardcoding decisions into software
 - (D) Storing and retrieving data efficiently
62. The main types of machine learning are:
- (A) Regression and Clustering
 - (B) Supervised, Unsupervised, and Reinforcement
 - (C) Linear and Non-linear
 - (D) Classification and Prediction
63. Regression problems are used for:
- (A) Predicting continuous values
 - (B) Predicting categorical values
 - (C) Grouping similar data
 - (D) Detecting anomalies
64. Classification problems are used for:
- (A) Predicting age of a person
 - (B) Predicting if an email is spam or not
 - (C) Predicting stock price
 - (D) Estimating rainfall amount

65. Supervised learning requires:
- (A) Labeled data
 - (B) Unlabeled data
 - (C) No data at all
 - (D) Only Reinforcement signals
66. Unsupervised learning is mainly used for:
- (A) Regression tasks
 - (B) Clustering data without labels
 - (C) Classification tasks
 - (D) Forecasting trends
67. Which of the following is an example of supervised learning?
- (A) Market basket analysis
 - (B) Spam email detection
 - (C) Customer segmentation
 - (D) Anomaly detection
68. Which of the following is an example of unsupervised learning?
- (A) Predicting house prices
 - (B) Classifying animals as dogs or cats
 - (C) Clustering customers based on purchases
 - (D) Detecting fraud with past labels
69. In machine learning, a feature refers to:
- (A) A dataset
 - (B) An individual measurable property of data
 - (C) The final prediction
 - (D) The training algorithm
70. Which machine learning task predicts values like sales, temperature, or weight?
- (A) Classification
 - (B) Regression
 - (C) Clustering
 - (D) Dimensionality reduction
71. Overfitting occurs when:
- (A) The model performs well on training but poorly on test data
 - (B) The model performs well on both training and test data
 - (C) The model performs poorly on both training and test data
 - (D) The model ignores training data
72. The opposite of overfitting is:
- (A) Clustering
 - (B) Underfitting
 - (C) Classification
 - (D) Regression

73. Which of the following is not a supervised learning algorithm?
- (A) Linear Regression
 - (B) Decision Trees
 - (C) K-Means Clustering
 - (D) Support Vector Machines
74. Which algorithm is most suitable for grouping customers based on behavior?
- (A) Linear Regression
 - (B) Logistic Regression
 - (C) K-Means Clustering
 - (D) Naive Bayes
75. Which of the following is an evaluation metric for classification problems?
- (A) Accuracy
 - (B) Mean Absolute Error
 - (C) Root Mean Square Error
 - (D) Adjusted
76. Reinforcement Learning uses:
- (A) Rewards and penalties
 - (B) Labeled data
 - (C) Clustering
 - (D) Regression
77. In supervised learning, the dataset is usually divided into:
- (A) Training and Test sets
 - (B) Validation and Feedback sets
 - (C) Training and Clustering sets
 - (D) Feature and Attribute sets
78. An example of classification is:
- (A) Predicting house rent
 - (B) Predicting tomorrow's temperature
 - (C) Determining if a tumor is malignant or benign
 - (D) Estimating employee salary
79. In ML, a "label" refers to:
- (A) The input variable
 - (B) The output variable we want to predict
 - (C) The preprocessing step
 - (D) A data collection method
80. Which learning approach is most suitable for playing chess through trial and error?
- (A) Supervised Learning
 - (B) Unsupervised Learning
 - (C) Reinforcement Learning
 - (D) Semi-supervised Learning

81. Neural Networks are inspired by:
- (A) Circuits in computers
 - (B) Biological neurons in the brain
 - (C) Chemical reactions
 - (D) Mechanical systems
82. Which is the basic unit of an artificial neural network?
- (A) Neuron
 - (B) Layer
 - (C) Synapse
 - (D) Weight
83. In biological neurons, information passes through :
- (A) Axons and dendrites
 - (B) Chips and processors
 - (C) Tables and charts
 - (D) Layers and weights
84. The function of weights in a neural network is to:
- (A) Store raw data
 - (B) Determine the strength of connections
 - (C) Reduce dimensionality
 - (D) Remove noise
85. Which neural network model is commonly used for image recognition?
- (A) Convolutional Neural Network (CNN)
 - (B) Recurrent Neural Network (RNN)
 - (C) Perceptron
 - (D) Hopfield Network
86. The first model of an artificial neuron was called:
- (A) Backpropagation model
 - (B) Perceptron
 - (C) Hebbian model
 - (D) Boltzmann machine
87. An activation function in neural networks is used to:
- (A) Normalize data
 - (B) Introduce non-linearity
 - (C) Remove duplicates
 - (D) Reduce features
88. Which activation function outputs values between 0 and 1?
- (A) ReLU
 - (B) Sigmoid
 - (C) Tanh
 - (D) Linear
89. The Tanh activation function outputs values between:
- (A) 0 and 1
 - (B) -1 and +1
 - (C) 0 and ∞
 - (D) $-\infty$ and $+\infty$
90. Which algorithm is commonly used to train neural networks?
- (A) Decision Trees
 - (B) Backpropagation
 - (C) K-Means
 - (D) Linear Regression

91. A feedforward neural network means:
 - (A) Data moves only in one direction
 - (B) Data moves in both directions
 - (C) Data moves randomly
 - (D) Data forms loops
92. Recurrent Neural Networks (RNNs) are especially useful for:
 - (A) Image classification
 - (B) Sequence and time-series data
 - (C) Static data
 - (D) Dimensionality reduction
93. The main challenge with training deep neural networks is:
 - (A) Too many labels
 - (B) Vanishing and exploding gradients
 - (C) Lack of features
 - (D) Limited activation functions
94. Dropout in neural networks is a technique used for:
 - (A) Reducing overfitting
 - (B) Increasing learning rate
 - (C) Removing irrelevant features
 - (D) Scaling data
95. In neural networks, bias helps to:
 - (A) Shift the activation function
 - (B) Remove redundant features
 - (C) Increase dataset size
 - (D) Perform clustering
96. Which type of network allows feedback loops?
 - (A) Feedforward Network
 - (B) Recurrent Neural Network
 - (C) Convolutional Network
 - (D) Multi-layer Perceptron
97. A deep neural network usually refers to a network with:
 - (A) One hidden layer
 - (B) More than one hidden layer
 - (C) Only input and output layers
 - (D) No activation function
98. Which neural network model is energy-based?
 - (A) CNN
 - (B) RNN
 - (C) Hopfield Network
 - (D) Perceptron
99. Artificial Intelligence and Neural Networks are connected because:
 - (A) Neural networks are a subset of AI techniques
 - (B) AI is a subset of neural networks
 - (C) Both are unrelated
 - (D) Both only deal with statistics
100. The main advantage of neural networks is:
 - (A) Ability to learn complex non-linear relationships
 - (B) Simplicity of interpretation
 - (C) Requirement of no data
 - (D) Manual feature engineering

Rough Work

Example :

Question :

Q.1 (A) ● (C) (D)

Q.2 (A) (B) ● (D)

Q.3 (A) ● (C) (D)

5. Each question carries equal marks. Marks will be awarded according to the number of correct answers you have.
6. All answers are to be given on OMR Answer Sheet only. Answers given anywhere other than the place specified in the answer sheet will not be considered valid.
7. Before writing anything on the OMR Answer Sheet, all the instructions given in it should be read carefully.
8. After the completion of the examination, candidates should leave the examination hall only after providing their OMR Answer Sheet to the invigilator. Candidate can carry their Question Booklet.
9. There will be no negative marking.
10. Rough work, if any, should be done on the blank pages provided for the purpose in the booklet.
11. To bring and use of log-book, calculator, pager & cellular phone in examination hall is prohibited.
12. In case of any difference found in English and Hindi version of the question, the English version of the question will be held authentic.

Impt. On opening the question booklet, first check that all the pages of the question booklet are printed properly. If there is any discrepancy in the question Booklet, then after showing it to the invigilator, get another question Booklet of the same series.

उदाहरण :

प्रश्न :

प्रश्न 1 (A) ● (C) (D)

प्रश्न 2 (A) (B) ● (D)

प्रश्न 3 (A) ● (C) (D)

5. प्रत्येक प्रश्न के अंक समान हैं। आपके जितने उत्तर सही होंगे, उन्हीं के अनुसार अंक प्रदान किये जायेंगे।
6. सभी उत्तर केवल ओ०एम०आर० उत्तर-पत्रक (OMR Answer Sheet) पर ही दिये जाने हैं। उत्तर-पत्रक में निर्धारित स्थान के अलावा अन्यत्र कहीं पर दिया गया उत्तर मान्य नहीं होगा।
7. ओ०एम०आर० उत्तर-पत्रक (OMR Answer Sheet) पर कुछ भी लिखने से पूर्व उसमें दिये गये सभी अनुदेशों को सावधानीपूर्वक पढ़ लिया जाये।
8. परीक्षा समाप्ति के उपरान्त परीक्षार्थी कक्ष निरीक्षक को अपनी OMR Answer Sheet उपलब्ध कराने के बाद ही परीक्षा कक्ष से प्रस्थान करें। परीक्षार्थी अपने साथ प्रश्न-पुस्तिका ले जा सकते हैं।
9. निगेटिव मार्किंग नहीं है।
10. कोई भी रफ कार्य, प्रश्न-पुस्तिका में, रफ-कार्य के लिए दिए खाली पेज पर ही किया जाना चाहिए।
11. परीक्षा-कक्ष में लॉग-बुक, कैल्कुलेटर, पेजर तथा सेल्युलर फोन ले जाना तथा उसका उपयोग करना वर्जित है।
12. प्रश्न के हिन्दी एवं अंग्रेजी रूपान्तरण में भिन्नता होने की दशा में प्रश्न का अंग्रेजी रूपान्तरण ही मान्य होगा।

महत्वपूर्ण: प्रश्नपुस्तिका खोलने पर प्रथमतः जाँच कर देख लें कि प्रश्नपुस्तिका के सभी पृष्ठ भलीभाँति छपे हुए हैं। यदि प्रश्नपुस्तिका में कोई कमी हो, तो कक्षनिरीक्षक को दिखाकर उसी सिरिज की दूसरी प्रश्नपुस्तिका प्राप्त कर लें।