CBCS SCHEME - Make-Up Exam

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

BCV601

Sixth Semester B.E/B.Tech. Degree Examination, June/July 2025
Design of RCC Structures

Time: 3 hrs. Max. Marks:100

Note: *1. Answer any FIVE full questions, choosing ONE full question from each module.*
2. M : Marks , L: Bloom's level , C: Course outcomes.
3. Assume suitable additional data, if necessary.
4. Use of IS456-2000, SP-16 is permitted.

		Module – 1	M	L	C
1	a.	Compare working Stress Method and Limit State Method.	10	L2	CO1
	b.	Explain : (i) Characteristic strength. (ii) Design strength (iii) Characteristic load (iv) Design load	10	L2	CO1
		OR			
2	a.	Explain the terms : Balanced, Under reinforced and over reinforced sections in Limit State method.	6	L2	CO1
	b.	A rectangular section beam 200 mm wide by 450 mm overall depth is reinforced with 3 bars of 16 mm diameter at an effective depth of 420 mm. Two hanger bars of 12 mm diameter are provided at the compression face. The effective span of the beam is 5 m. The beam supports a service live load of 10 kN/m. If f_{ck} = 20 N/mm^2 and f_y = 415 N/mm^2. Compute (i) Short Term Deflection (ii) Long Term Deflections according to IS code specifications.	14	L3	CO1
		Module – 2			
3	a.	Prove $x_u = 0.48\,d$ from properties of Balanced section of steel grade – Fe-415.	5	L5	CO2

	b.	A Rectangular beam of size 200 mm width, 450 mm effective depth is reinforced with 3 Nos of 16 mm diameter bars. Interpret the following cases : Case (i) : Safe moment of Resistance of the section. Case (ii) : Safe moment of Resistance of the section, if steel increases to 3 Nos of 20 mm diameter bars. Case (iii) : Strain in steel for both cases (i) and (ii). Use M-20 grade of concrete and Fe-415 grade of steel.	**15**	**L5**	**CO2**
		OR			
4	a.	What are the situations where doubly reinforced beams are used in structural design?	**4**	**L1**	**CO2**
	b.	Determine the ultimate Flexural strength of T-beam for the following section properties : Width of flange = 800 mm Width of Rib = 300 mm Depth of Flange = 150 mm Effective depth = 420 mm Area of steel = 1470 mm^2 M-25 grade of concrete and Fe-415 grade of steel.	**16**	**L5**	**CO2**
		Module – 3			
5		Design a singly reinforced beam of clear span 5 m to support a working live load 12 kN/m. Adopt M-20 grade of concrete and Fe-415 grade of HYSD bars. Sketch the reinforcement details.	**20**	**L5**	**CO3**
		OR			
6		Design a Doubly reinforced concrete beam of rectangular section using the following data : Effective span = 8 m Working live load = 30 kN/m Overall depth restricted to 650 mm Width of beam = 300 mm M-20 grade of concrete and Fe-415 grade of steel. Sketch the reinforcement details.	**20**	**L5**	**CO3**

		Module – 4			
7		Design a RCC slab for a room size 4m × 5m. The slab is supported all around on wall of width 300 mm. The slab has to carry a live load of 4 kN/m^2 and floor finish 1 KN/m^2. Use M-20 grade of concrete and Fe-415 grade of steel. Assume corners are held down. Sketch the reinforcement details.	**20**	**L5**	**CO4**
		OR			
8		Design one of the flights of stairs of a school building spanning between landing beams to suit the following data : Type of stair case-Dog-legged Number of steps in flight = 12 Tread = T = 300 mm Riser = R = 160 mm Width of landing beams = 400 mm Materials M20 grade of concrete and Fe-415 grade of steel. Also show the reinforcement details.	**20**	**L5**	**CO4**
		Module – 5			
9		A short column located at the corner of a building is subjected to an factored axial load of 2000 KN together with factored moments of 75 kN-m and 60 kN-m acting in perpendicular planes. The size of column is fixed as 450 mm×450 mm. Adopting M-20 grade of concrete and Fe-415 grade of steel. Design suitable reinforcements in column section.	**20**	**L3**	**CO5**
		OR			
10		A RCC column of size 400 mm × 400 mm supports an axial service load of 1000 kN. The safe bearing capacity of soil at site is 200 kN/m^2. Adopting M-20 grade of concrete and Fe-415 grade of steel. Design suitable footing for the column and sketch the details of reinforcements.	**20**	**L3**	**CO5**

* * * * *

CBCS SCHEME - Make-Up Exam

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

BCV602

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Irrigation Engineering and Hydraulic Structures

Time: 3 hrs. Max. Marks: 100

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. M : Marks , L: Bloom's level , C: Course outcomes.

		Module – 1	M	L	C
Q.1	a.	Define the term irrigation and what are the types of flow irrigation? Explain any two flow irrigation system.	6	L2	CO1
	b.	Define duty, delta and Base period. Derive an expression to establish relation between them.	6	L2	CO1
	c.	After how many days will you supply water to soil in order to ensure sufficient irrigation of the given crop if i) Field capacity of soil = 28% ii) Permanent wilting point = 13% iii) Effective depth of root zone = 70 cm iv) Dry density of soil = 1.3 gm/cc v) Daily consumptive use of water of given crop = 12 mm	8	L3	CO1
		OR			
Q.2	a.	List the benefits and ill effects of irrigation.	6	L1	CO1
	b.	List and explain Irrigation efficiencies.	6	L2	CO1
	c.	An irrigation canal has gross command area of 80000 hectors out of which 85% is culturable. The intensity of irrigation for Kharif season is 30% and for Rabi season 60%. Find the discharge required at the head of the canal if the duty at its head is 800 ho/cu for Kharif season and 1700 ho/cu for rabi season.	8	L3	CO1
		Module – 2			
Q.3	a.	Explain various considerations for canal alignment.	8	L2	CO1
	b.	Design an irrigation channel to carry a discharge of 45 cumecs. Assume N = 0.0225 and M = 1. The channel has a bed slope of 0.16 meter per kilometer. Use Kennedys theory and trial depth D as 1.8 m.	12	L3	CO1
		OR			
Q.4	a.	Explain with neat sketch the storage zones of a reservoir.	8	L2	CO1
	b.	The monthly yield of water from a catchment is given below. Determine the minimum capacity of the reservoir by mass curve method if the flow is drawn at a uniform rate. Values are given in million cubic meters	12	L3	CO1

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Inflow volume	1.4	2.1	2.8	8.4	11.9	11.9	7.7	2.8	2.25	2.24	1.96	1.68

		Module – 3			
Q.5	**a.**	Discuss briefly forces acting on gravity dam with the help of a neat sketch.	**10**	**L2**	**CO1**
	b.	Design the practical profile of a gravity dam made of stone masonry given the following data: RL of base of dam = 198 m RL of HFL of reservoir = 228 m Specific gravity of masonry = 2.4 Safe compressive stress in = 1200 kN/m^2 masonry	**10**	**L3**	**CO1**
		OR			
Q.6	**a.**	Discuss in brief various modes of failure of gravity dam.	**6**	**L2**	**CO1**
	b.	Explain step by step graphical procedure to be adopted for analyzing the stability of gravity dam.	**8**	**L2**	**CO1**
	c.	Design and draw the practical profile of a gravity dam to a suitable scale, when height of water to be stored = 55 m, specific gravity of concrete = 2.4, free board = 2.75 m	**6**	**L3**	**CO1**
		Module – 4			
Q.7	**a.**	Explain with neat sketches different types of earth dams.	**10**	**L2**	**CO2**
	b.	Explain the causes of failure of earth dam.	**10**	**L2**	**CO2**
		OR			
Q.8	**a.**	Define a spillway. Write neat sketches of different types of spillways.	**10**	**L2**	**CO2**
	b.	Describe the design principles that are involved in the design of ogee spillway.	**10**	**L2**	**CO2**
		Module – 5			
Q.9	**a.**	Draw a neat sketch of diversion head works and indicate various components of the system. Briefly indicate the function of each component.	**10**	**L2**	**CO3**
	b.	Briefly outline Khosla's theory on the design of weirs on permeable foundation. Enumerate the various corrections that are needed in the application of this theory.	**10**	**L2**	**CO3**
		OR			
Q.10	**a.**	Define a Weir and Barrage with the help of a neat sketch.	**6**	**L2**	**CO3**
	b.	Explain Bligh's creep theory for the design of impervious floor weir.	**6**	**L2**	**CO3**
	c.	Briefly explain silt ejectors and silt excluders.	**8**	**L2**	**CO3**

* * * * *

CBCS SCHEME

BCV657C

USN | | | | | | | | | |

Question Paper Version : A

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Data Analytics for Civil Engineers

Time: 1 hr.] [Max. Marks: 50

INSTRUCTIONS TO THE CANDIDATES

1. Answer all the **fifty** questions, each question carries one mark.
2. Use only **Black ball point pen** for writing / darkening the circles.
3. **For each question, after selecting your answer, darken the appropriate circle corresponding to the same question number on the OMR sheet.**
4. Darkening two circles for the same question makes the answer invalid.
5. **Damaging/overwriting, using whiteners** on the **OMR** sheets are strictly prohibited.

---

1. Which of the following frameworks is commonly used in data mining and knowledge?
   a) SEMMA b) CRISP - DM c) Both (a) & (b) d) None of these

2. Which of the following best describes the relationship between knowledge and data?
   a) Data is raw facts, while knowledge is derived from interpreting data
   b) Knowledge and data are the same and can be used interchangeably
   c) Data is processed information , whereas knowledge is unstructured data.
   d) Knowledge is collected first and details generated from it.

3. Which of the following correctly represents the hierarchy from raw data to knowledge?
   a) Knowledge → Information → Data b) Data → Knowledge → Information
   c) Data → Information → Knowledge d) Information → Data → Knowledge

4. Which of the following is NOT a measure of central tendency in descriptive statistics?
   a) Mean b) Median c) Mode d) Standard deviation

5. Which of the following is NOT a common criterion for assessing knowledge?
   a) Accuracy b) Relevance c) Randomness d) Consistency

6. Which of the following is a key criterion for evaluating the quality of knowledge?
   a) Validity b) Ambiguity c) Inconsistency d) Irrelevance

7. Which of the following is the main purpose of inferential statistics?
   a) To summarize and describe data
   b) To make predications or generalizations about a population based on a sample
   c) To organize raw data into labels and graphs
   d) To calculate the exact values for an entire population.

8. What is the first step in the data analysis process?
   a) Data cleaning b) Data collection
   c) Data visualization d) Data interpretation

9. What does the 'S' in SEMMA stand for?
   a) Sample b) Structure c) Select d) Survey

10. What does CRISP – DM stand for?
    a) Cross – Industry standard process for Data Mining
    b) Common Research Information for Statical Models
    c) Critical Reporting and Information for Statical Models
    d) Custom Research and Information for Data Management

11. Which of the following methods is used to handle missing values in the data set during attribute understanding?
    a) Featuring scale b) Imputation
    c) One hot coding d) Correlation analysis

12. Which type of data transformation is typically required for numerical attributes in machine learning models?
    a) One – hot encoding b) Normalization
    c) Label encoding d) Feature extraction

13. Which of the following is an example of a categorical attribute?
    a) Age b) Income c) Gender d) Temperature

14. One – dimensional (1D) data is typically represented as :
    a) A matrix with multiple rows and columns
    b) A single column or row of values
    c) A cube with depth, width and height
    d) A set of attributes with values linked by relationship

15. One – dimensional (1D) data represents :
    a) A sequence of values from multiple variables
    b) A sequence of values from single variables
    c) Multiple rows and columns of data
    d) A multi dimensional array

16. Which of the following is a measure of the shape of a distribution?
    a) Mean b) Skewness c) Standard deviation d) Variance

17. A data set with a positive skewness indicates that
    a) Most of the data are concentrated on the right side
    b) Most of the data are concentrated on the left side
    c) The data is symmetrical
    d) The data is bimodal

18. Which of the following would likely be true for a perfectly symmetrical distribution?
    a) The mean less than the median
    b) The skewness is 0
    c) The data is positively skewed
    d) The mode is greater than the mean

19. Which of the following is a method to standardize multi dimensional data, especially when the variables are having different scales?
a) Normalization b) K – mean clustering
c) Decision tress d) Variance analysis

20. Which of the following is a common method for measuring the spread or variability in multidimensional data?
a) Covariance Matrix b) Mean Absolute Error (MAE)
c) Standard error of the mean d) Mode

21. Which of the following is an example of a model class?
a) Linear Regression b) Support vector machine
c) Decision tree d) All of these

22. Which of the following is an example of a black – box model?
a) Logistic regression b) Decision tree
c) Random forest d) Linear regression

23. Fitting criteria are used to :
a) Choose the best training data set
b) Evaluate the quality of the model based on the test data
c) Optimize the parameters of the model to fit the data well
d) Select the features for the model

24. Which of the following is commonly used as a fitting criterion for regression problems?
a) Mean Squared Error (MSE) b) Gini Index
c) Entropy d) R - Squared

25. Which of the following is a common score function used for classification problems?
a) Accuracy b) Log - loss c) AUC - ROC d) All of these

26. Which of the following best describes a "Score function" in machine learning?
a) A function that evaluates the performance of a model
b) A function that predicts the target variable
c) A function used to optimize the model parameters
d) A function used to preprocess the data

27. Which of the following is a measure of interestingness in association rule mining?
a) Support b) Mean Squared Error (MSE)
c) R – Squared d) Normalized Mutual Information (NMI)

28. Which of the following is NOT a common measure of interestingness for association rules?
a) Confidence b) Lift c) R - Squared d) Support

29. What is the key advantage of using a closed form solution for model fitting?
a) It guarantees and optimal solution in less time
b) It avoids the need for an iterative process
c) It is suitable for very large data sets
d) It works well for highly non linear data

30. Which of the following is a closed form solution for fitting a linear model?
a) Stochastic Gradient Descent (SGD) b) Newton's method
c) The Normal equation d) Support vector machines

31. Which of the following best describes data selection in Machine learning?
a) Choosing relevant features from the data set
b) Deciding which data samples to include in the training set
c) Preprocessing the data by removing noise
d) Deciding on the evaluation metric to use

32. Which of the following is a common technique for selecting a subset of data for model training?
a) Cross validation b) K – means clustering
c) Principle component analysis d) Bagging

33. What is "Sampling Bias" in data selection?
a) The situation where the training data is not representative of the population
b) When the data is over sampling for specific features
c) The failure to normalize the data set
d) A technique used to reduce dimensionality

34. Which method can help in reducing selection bias when selecting data samples for training?
a) Random sampling b) Standardization
c) Data augmentation d) Feature scaling

35. Feature selection is the process of :
a) Identifying and removing irrelevant or redundant features from the data
b) Scaling features to a common range
c) Creating new features from the existing ones
d) Splitting the data into training and test sets

36. In pavement design, the use of the Elastic layered theory primarily focuses on :
a) Determining the maximum load that the pavement can bear
b) Predicting to life span of the pavement
c) Estimating the structural capacity of different pavement layers
d) Identifying traffic volume patterns

37. Which machine learning technique can be used to predict traffic flow patterns based on historical data in civil engineering?
a) Decision trees b) Artificial Neural Networks (ANNs)
c) Linear programming d) Time – series forecasting

38. Which method is typically used in structural health monitoring to detect changes in the behavior of a structure overtime?
a) Vibration based monitoring b) Cross validation
c) Image recognition d) particle swarm optimization

39. Which of the following is an example of a machine learning technique that can be used for predictive maintenance in civil engineering?
    a) Decision trees
    b) Random Forests
    c) Support Vector Machines (SVM)
    d) All of these

40. Which data analysis method is commonly employed to optimize the design of a bridge in terms of material usage, cost and local capacity?
    a) Multi – objective optimization
    b) Time series analysis
    c) Random sampling
    d) Hierarchical clustering

41. Which of the following is a characteristic of hierarchical clustering?
    A) It always requires the number of clusters to be predefined
    B) The dendrogram is used to visualize the hierarchical structure of clusters
    C) Hierarchical clustering is a non – iterative method
    D) Agglomerative hierarchical clustering begins by assuming all data points belong to a single clustering.
    a) (A) & (C) only
    b) (B) & (D) only
    c) (B), (C) & (D) only
    d) (A), (B) &(D) only

42. Which of the following statements best describes a Self Organization Map (SOM)?
    a) SOM is a type of supervised learning algorithm used for classification tasks.
    b) SOM reduces the dimensionality of data by mapping high dimensional input to a lower – dimensional grid while preserving topological relationship.
    c) SOM is based on the back propagation algorithm, similar to other neural network
    d) None of these

43. Which of the following statements about dissimilarity measures is correct?
    a) Dissimilarity measures are used to calculate how similar two objects are in clustering problems.
    b) The Euclidean distance is a common dissimilarity measure that calculates the square root of the sum of square difference between corresponding features.
    c) Dissimilarity measures are used to quantify the difference between objects in terms of a similarity metric
    d) None of these

44. What is the primary function of a Self Organizing Map (SOM)?
    a) Classification
    b) Regression
    c) Clustering and Visualization
    d) Anomaly detection

45. Which of the following best describes the learning process in a SOM?
    a) Supervised learning with labels
    b) Reinforcement learning
    c) Unsupervised learning using competitive learning
    d) Genetic algorithm based learning

46. In a SOM, which of the following happens during the training process.
    a) Only the Best Matching Unit (BMU) is updated
    b) The weights of all units are updated equally
    c) All neurons receive the same input signal
    d) None of these

47. Which of the following is true about the topological structure in SOMs?
    a) The topological structure is destored during training
    b) Similar data points map to distant neurons
    c) The map preserves the topological relationships between data points
    d) None of these

48. What does a neighborhood function in a SOM control?
    a) The number of neurons in the map
    b) The distance between the data points
    c) The influence of neighboring neurons on the weight update process
    d) The number of layers in the map

49. Which of the following is a common use of data analysis in Civil engineering?
    a) Traffic congestion forecasting
    b) Structural load analysis
    c) Analyzing tensile strength of materials
    d) All of these

50. In a structural engineering, which type of data would be most important for conducitng a finite element analysis?
    a) Soil Data
    b) Material properties , geometric dimensions and loading condition
    c) Weather data for construction scheduling
    d) None of these

* * * * *

CBCS SCHEME

BCV657C

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

Question Paper Version : B

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Data Analytics for Civil Engineers

Time: 1 hr.] [Max. Marks: 50

INSTRUCTIONS TO THE CANDIDATES

1. Answer all the **fifty** questions, each question carries one mark.
2. Use only **Black ball point pen** for writing / darkening the circles.
3. **For each question, after selecting your answer, darken the appropriate circle corresponding to the same question number on the OMR sheet.**
4. Darkening two circles for the same question makes the answer invalid.
5. **Damaging/overwriting, using whiteners** on the **OMR** sheets are strictly prohibited.

---

1. Which of the following best describes data selection in Machine learning?
   a) Choosing relevant features from the data set
   b) Deciding which data samples to include in the training set
   c) Preprocessing the data by removing noise
   d) Deciding on the evaluation metric to use

2. Which of the following is a common technique for selecting a subset of data for model training?
   a) Cross validation b) K – means clustering
   c) Principle component analysis d) Bagging

3. What is "Sampling Bias" in data selection?
   a) The situation where the training data is not representative of the population
   b) When the data is over sampling for specific features
   c) The failure to normalize the data set
   d) A technique used to reduce dimensionality

4. Which method can help in reducing selection bias when selecting data samples for training?
   a) Random sampling b) Standardization
   c) Data augmentation d) Feature scaling

5. Feature selection is the process of :
   a) Identifying and removing irrelevant or redundant features from the data
   b) Scaling features to a common range
   c) Creating new features from the existing ones
   d) Splitting the data into training and test sets

6. In pavement design, the use of the Elastic layered theory primarily focuses on :
   a) Determining the maximum load that the pavement can bear
   b) Predicting to life span of the pavement
   c) Estimating the structural capacity of different pavement layers
   d) Identifying traffic volume patterns

7. Which machine learning technique can be used to predict traffic flow patterns based on historical data in civil engineering?
   a) Decision trees b) Artificial Neural Networks (ANNs)
   c) Linear programming d) Time – series forecasting

8. Which method is typically used in structural health monitoring to detect changes in the behavior of a structure overtime?
   a) Vibration based monitoring b) Cross validation
   c) Image recognition d) particle swarm optimization

9. Which of the following is an example of a machine learning technique that can be used for predictive maintenance in civil engineering?
   a) Decision trees b) Random Forests
   c) Support Vector Machines (SVM) d) All of these

10. Which data analysis method is commonly employed to optimize the design of a bridge in terms of material usage, cost and local capacity?
    a) Multi – objective optimization b) Time series analysis
    c) Random sampling d) Hierarchical clustering

11. Which of the following frameworks is commonly used in data mining and knowledge?
    a) SEMMA b) CRISP - DM c) Both (a) & (b) d) None of these

12. Which of the following best describes the relationship between knowledge and data?
    a) Data is raw facts, while knowledge is derived from interpreting data
    b) Knowledge and data are the same and can be used interchangeably
    c) Data is processed information , whereas knowledge is unstructured data.
    d) Knowledge is collected first and details generated from it.

13. Which of the following correctly represents the hierarchy from raw data to knowledge?
    a) Knowledge → Information → Data b) Data → Knowledge → Information
    c) Data → Information → Knowledge d) Information → Data → Knowledge

14. Which of the following is NOT a measure of central tendency in descriptive statistics?
    a) Mean b) Median c) Mode d) Standard deviation

15. Which of the following is NOT a common criterion for assessing knowledge?
    a) Accuracy b) Relevance c) Randomness d) Consistency

16. Which of the following is a key criterion for evaluating the quality of knowledge?
    a) Validity b) Ambiguity c) Inconsistency d) Irrelevance

17. Which of the following is the main purpose of inferential statistics?
    a) To summarize and describe data
    b) To make predications or generalizations about a population based on a sample
    c) To organize raw data into labels and graphs
    d) To calculate the exact values for an entire population.

18. What is the first step in the data analysis process?
    a) Data cleaning
    b) Data collection
    c) Data visualization
    d) Data interpretation

19. What does the 'S' in SEMMA stand for?
    a) Sample
    b) Structure
    c) Select
    d) Survey

20. What does CRISP – DM stand for?
    a) Cross – Industry standard process for Data Mining
    b) Common Research Information for Statical Models
    c) Critical Reporting and Information for Statical Models
    d) Custom Research and Information for Data Management

21. Which of the following is a characteristic of hierarchical clustering?
    A) It always requires the number of clusters to be predefined
    B) The dendrogram is used to visualize the hierarchical structure of clusters
    C) Hierarchical clustering is a non – iterative method
    D) Agglomerative hierarchical clustering begins by assuming all data points belong to a single clustering.
    a) (A) & (C) only
    b) (B) & (D) only
    c) (B), (C) & (D) only
    d) (A), (B) &(D) only

22. Which of the following statements best describes a Self Organization Map (SOM)?
    a) SOM is a type of supervised learning algorithm used for classification tasks.
    b) SOM reduces the dimensionality of data by mapping high dimensional input to a lower – dimensional grid while preserving topological relationship.
    c) SOM is based on the back propagation algorithm, similar to other neural network
    d) None of these

23. Which of the following statements about dissimilarity measures is correct?
    a) Dissimilarity measures are used to calculate how similar two objects are in clustering problems.
    b) The Euclidean distance is a common dissimilarity measure that calculates the square root of the sum of square difference between corresponding features.
    c) Dissimilarity measures are used to quantify the difference between objects in terms of a similarity metric
    d) None of these

24. What is the primary function of a Self Organizing Map (SOM)?
    a) Classification
    b) Regression
    c) Clustering and Visualization
    d) Anomaly detection

25. Which of the following best describes the learning process in a SOM?
    a) Supervised learning with labels
    b) Reinforcement learning
    c) Unsupervised learning using competitive learning
    d) Genetic algorithm based learning

26. In a SOM, which of the following happens during the training process.
    a) Only the Best Matching Unit (BMU) is updated
    b) The weights of all units are updated equally
    c) All neurons receive the same input signal
    d) None of these

27. Which of the following is true about the topological structure in SOMs?
    a) The topological structure is destored during training
    b) Similar data points map to distant neurons
    c) The map preserves the topological relationships between data points
    d) None of these

28. What does a neighborhood function in a SOM control?
    a) The number of neurons in the map
    b) The distance between the data points
    c) The influence of neighboring neurons on the weight update process
    d) The number of layers in the map

29. Which of the following is a common use of data analysis in Civil engineering?
    a) Traffic congestion forecasting
    b) Structural load analysis
    c) Analyzing tensile strength of materials
    d) All of these

30. In a structural engineering, which type of data would be most important for conducitng a finite element analysis?
    a) Soil Data
    b) Material properties , geometric dimensions and loading condition
    c) Weather data for construction scheduling
    d) None of these

31. Which of the following is an example of a model class?
    a) Linear Regression
    b) Support vector machine
    c) Decision tree
    d) All of these

32. Which of the following is an example of a black – box model?
    a) Logistic regression
    b) Decision tree
    c) Random forest
    d) Linear regression

33. Fitting criteria are used to :
    a) Choose the best training data set
    b) Evaluate the quality of the model based on the test data
    c) Optimize the parameters of the model to fit the data well
    d) Select the features for the model

34. Which of the following is commonly used as a fitting criterion for regression problems?
a) Mean Squared Error (MSE) b) Gini Index
c) Entropy d) R - Squared

35. Which of the following is a common score function used for classification problems?
a) Accuracy b) Log - loss c) AUC - ROC d) All of these

36. Which of the following best describes a "Score function" in machine learning?
a) A function that evaluates the performance of a model
b) A function that predicts the target variable
c) A function used to optimize the model parameters
d) A function used to preprocess the data

37. Which of the following is a measure of interestingness in association rule mining?
a) Support b) Mean Squared Error (MSE)
c) R – Squared d) Normalized Mutual Information (NMI)

38. Which of the following is NOT a common measure of interestingness for association rules?
a) Confidence b) Lift c) R - Squared d) Support

39. What is the key advantage of using a closed form solution for model fitting?
a) It guarantees and optimal solution in less time
b) It avoids the need for an iterative process
c) It is suitable for very large data sets
d) It works well for highly non linear data

40. Which of the following is a closed form solution for fitting a linear model?
a) Stochastic Gradient Descent (SGD) b) Newton's method
c) The Normal equation d) Support vector machines

41. Which of the following methods is used to handle missing values in the data set during attribute understanding?
a) Featuring scale b) Imputation
c) One hot coding d) Correlation analysis

42. Which type of data transformation is typically required for numerical attributes in machine learning models?
a) One – hot encoding b) Normalization
c) Label encoding d) Feature extraction

43. Which of the following is an example of a categorical attribute?
a) Age b) Income c) Gender d) Temperature

44. One – dimensional (1D) data is typically represented as :
a) A matrix with multiple rows and columns
b) A single column or row of values
c) A cube with depth, width and height
d) A set of attributes with values linked by relationship

45. One – dimensional (1D) data represents :
a) A sequence of values from multiple variables
b) A sequence of values from single variables
c) Multiple rows and columns of data
d) A multi dimensional array

46. Which of the following is a measure of the shape of a distribution?
a) Mean b) Skewness c) Standard deviation d) Variance

47. A data set with a positive skewness indicates that
a) Most of the data are concentrated on the right side
b) Most of the data are concentrated on the left side
c) The data is symmetrical
d) The data is bimodal

48. Which of the following would likely be true for a perfectly symmetrical distribution?
a) The mean less than the median
b) The skewness is 0
c) The data is positively skewed
d) The mode is greater than the mean

49. Which of the following is a method to standardize multi dimensional data, especially when the variables are having different scales?
a) Normalization b) K – mean clustering
c) Decision tress d) Variance analysis

50. Which of the following is a common method for measuring the spread or variability in multidimensional data?
a) Covariance Matrix b) Mean Absolute Error (MAE)
c) Standard error of the mean d) Mode

*****

CBCS SCHEME

BCV657C

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

Question Paper Version : C

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Data Analytics for Civil Engineers

Time: 1 hr.] [Max. Marks: 50

INSTRUCTIONS TO THE CANDIDATES

1. Answer all the **fifty** questions, each question carries one mark.
2. Use only **Black ball point pen** for writing / darkening the circles.
3. **For each question, after selecting your answer, darken the appropriate circle corresponding to the same question number on the OMR sheet.**
4. Darkening two circles for the same question makes the answer invalid.
5. **Damaging/overwriting, using whiteners** on the OMR sheets are strictly prohibited.

---

1. Which of the following is an example of a model class?
   a) Linear Regression b) Support vector machine
   c) Decision tree d) All of these
2. Which of the following is an example of a black – box model?
   a) Logistic regression b) Decision tree
   c) Random forest d) Linear regression
3. Fitting criteria are used to :
   a) Choose the best training data set
   b) Evaluate the quality of the model based on the test data
   c) Optimize the parameters of the model to fit the data well
   d) Select the features for the model
4. Which of the following is commonly used as a fitting criterion for regression problems?
   a) Mean Squared Error (MSE) b) Gini Index
   c) Entropy d) R - Squared
5. Which of the following is a common score function used for classification problems?
   a) Accuracy b) Log - loss c) AUC - ROC d) All of these
6. Which of the following best describes a "Score function" in machine learning?
   a) A function that evaluates the performance of a model
   b) A function that predicts the target variable
   c) A function used to optimize the model parameters
   d) A function used to preprocess the data

7. Which of the following is a measure of interestingness in association rule mining?
   a) Support b) Mean Squared Error (MSE)
   c) R – Squared d) Normalized Mutual Information (NMI)
8. Which of the following is NOT a common measure of interestingness for association rules?
   a) Confidence b) Lift c) R - Squared d) Support
9. What is the key advantage of using a closed form solution for model fitting?
   a) It guarantees and optimal solution in less time
   b) It avoids the need for an iterative process
   c) It is suitable for very large data sets
   d) It works well for highly non linear data
10. Which of the following is a closed form solution for fitting a linear model?
    a) Stochastic Gradient Descent (SGD) b) Newton's method
    c) The Normal equation d) Support vector machines
11. Which of the following is a characteristic of hierarchical clustering?
    A) It always requires the number of clusters to be predefined
    B) The dendrogram is used to visualize the hierarchical structure of clusters
    C) Hierarchical clustering is a non – iterative method
    D) Agglomerative hierarchical clustering begins by assuming all data points belong to a single clustering.
    a) (A) & (C) only b) (B) & (D) only
    c) (B), (C) & (D) only d) (A), (B) &(D) only
12. Which of the following statements best describes a Self Organization Map (SOM)?
    a) SOM is a type of supervised learning algorithm used for classification tasks.
    b) SOM reduces the dimensionality of data by mapping high dimensional input to a lower – dimensional grid while preserving topological relationship.
    c) SOM is based on the back propagation algorithm, similar to other neural network
    d) None of these
13. Which of the following statements about dissimilarity measures is correct?
    a) Dissimilarity measures are used to calculate how similar two objects are in clustering problems.
    b) The Euclidean distance is a common dissimilarity measure that calculates the square root of the sum of square difference between corresponding features.
    c) Dissimilarity measures are used to quantify the difference between objects in terms of a similarity metric
    d) None of these
14. What is the primary function of a Self Organizing Map (SOM)?
    a) Classification b) Regression
    c) Clustering and Visualization d) Anomaly detection

15. Which of the following best describes the learning process in a SOM?
    a) Supervised learning with labels
    b) Reinforcement learning
    c) Unsupervised learning using competitive learning
    d) Genetic algorithm based learning

16. In a SOM, which of the following happens during the training process.
    a) Only the Best Matching Unit (BMU) is updated
    b) The weights of all units are updated equally
    c) All neurons receive the same input signal
    d) None of these

17. Which of the following is true about the topological structure in SOMs?
    a) The topological structure is destored during training
    b) Similar data points map to distant neurons
    c) The map preserves the topological relationships between data points
    d) None of these

18. What does a neighborhood function in a SOM control?
    a) The number of neurons in the map
    b) The distance between the data points
    c) The influence of neighboring neurons on the weight update process
    d) The number of layers in the map

19. Which of the following is a common use of data analysis in Civil engineering?
    a) Traffic congestion forecasting b) Structural load analysis
    c) Analyzing tensile strength of materials d) All of these

20. In a structural engineering, which type of data would be most important for conduciting a finite element analysis?
    a) Soil Data
    b) Material properties , geometric dimensions and loading condition
    c) Weather data for construction scheduling
    d) None of these

21. Which of the following best describes data selection in Machine learning?
    a) Choosing relevant features from the data set
    b) Deciding which data samples to include in the training set
    c) Preprocessing the data by removing noise
    d) Deciding on the evaluation metric to use

22. Which of the following is a common technique for selecting a subset of data for model training?
    a) Cross validation b) K – means clustering
    c) Principle component analysis d) Bagging

23. What is "Sampling Bias" in data selection?
    a) The situation where the training data is not representative of the population
    b) When the data is over sampling for specific features
    c) The failure to normalize the data set
    d) A technique used to reduce dimensionality

24. Which method can help in reducing selection bias when selecting data samples for training?
    a) Random sampling b) Standardization
    c) Data augmentation d) Feature scaling

25. Feature selection is the process of :
    a) Identifying and removing irrelevant or redundant features from the data
    b) Scaling features to a common range
    c) Creating new features from the existing ones
    d) Splitting the data into training and test sets

26. In pavement design, the use of the Elastic layered theory primarily focuses on :
    a) Determining the maximum load that the pavement can bear
    b) Predicting to life span of the pavement
    c) Estimating the structural capacity of different pavement layers
    d) Identifying traffic volume patterns

27. Which machine learning technique can be used to predict traffic flow patterns based on historical data in civil engineering?
    a) Decision trees b) Artificial Neural Networks (ANNs)
    c) Linear programming d) Time – series forecasting

28. Which method is typically used in structural health monitoring to detect changes in the behavior of a structure overtime?
    a) Vibration based monitoring b) Cross validation
    c) Image recognition d) particle swarm optimization

29. Which of the following is an example of a machine learning technique that can be used for predictive maintenance in civil engineering?
    a) Decision trees b) Random Forests
    c) Support Vector Machines (SVM) d) All of these

30. Which data analysis method is commonly employed to optimize the design of a bridge in terms of material usage, cost and local capacity?
    a) Multi – objective optimization b) Time series analysis
    c) Random sampling d) Hierarchical clustering

31. Which of the following methods is used to handle missing values in the data set during attribute understanding?
    a) Featuring scale b) Imputation
    c) One hot coding d) Correlation analysis

32. Which type of data transformation is typically required for numerical attributes in machine learning models?
    a) One – hot encoding b) Normalization
    c) Label encoding d) Feature extraction

33. Which of the following is an example of a categorical attribute?
    a) Age b) Income c) Gender d) Temperature

34. One – dimensional (1D) data is typically represented as :
a) A matrix with multiple rows and columns
b) A single column or row of values
c) A cube with depth, width and height
d) A set of attributes with values linked by relationship

35. One – dimensional (1D) data represents :
a) A sequence of values from multiple variables
b) A sequence of values from single variables
c) Multiple rows and columns of data
d) A multi dimensional array

36. Which of the following is a measure of the shape of a distribution?
a) Mean b) Skewness c) Standard deviation d) Variance

37. A data set with a positive skewness indicates that
a) Most of the data are concentrated on the right side
b) Most of the data are concentrated on the left side
c) The data is symmetrical
d) The data is bimodal

38. Which of the following would likely be true for a perfectly symmetrical distribution?
a) The mean less than the median
b) The skewness is 0
c) The data is positively skewed
d) The mode is greater than the mean

39. Which of the following is a method to standardize multi dimensional data, especially when the variables are having different scales?
a) Normalization b) K – mean clustering
c) Decision tress d) Variance analysis

40. Which of the following is a common method for measuring the spread or variability in multidimensional data?
a) Covariance Matrix b) Mean Absolute Error (MAE)
c) Standard error of the mean d) Mode

41. Which of the following frameworks is commonly used in data mining and knowledge?
a) SEMMA b) CRISP - DM c) Both (a) & (b) d) None of these

42. Which of the following best describes the relationship between knowledge and data?
a) Data is raw facts, while knowledge is derived from interpreting data
b) Knowledge and data are the same and can be used interchangeably
c) Data is processed information , whereas knowledge is unstructured data.
d) Knowledge is collected first and details generated from it.

43. Which of the following correctly represents the hierarchy from raw data to knowledge?
a) Knowledge → Information → Data b) Data → Knowledge → Information
c) Data → Information → Knowledge d) Information → Data → Knowledge

44. Which of the following is NOT a measure of central tendency in descriptive statistics?
a) Mean b) Median c) Mode d) Standard deviation

45. Which of the following is NOT a common criterion for assessing knowledge?
a) Accuracy b) Relevance c) Randomness d) Consistency

46. Which of the following is a key criterion for evaluating the quality of knowledge?
a) Validity b) Ambiguity c) Inconsistency d) Irrelevance

47. Which of the following is the main purpose of inferential statistics?
a) To summarize and describe data
b) To make predications or generalizations about a population based on a sample
c) To organize raw data into labels and graphs
d) To calculate the exact values for an entire population.

48. What is the first step in the data analysis process?
a) Data cleaning b) Data collection
c) Data visualization d) Data interpretation

49. What does the 'S' in SEMMA stand for?
a) Sample b) Structure c) Select d) Survey

50. What does CRISP – DM stand for?
a) Cross – Industry standard process for Data Mining
b) Common Research Information for Statical Models
c) Critical Reporting and Information for Statical Models
d) Custom Research and Information for Data Management

* * * * *

CBCS SCHEME

BCV657C

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

Question Paper Version : D

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Data Analytics for Civil Engineers

Time: 1 hr.] [Max. Marks: 50

INSTRUCTIONS TO THE CANDIDATES

1. Answer all the **fifty** questions, each question carries one mark.
2. Use only **Black ball point pen** for writing / darkening the circles.
3. **For each question, after selecting your answer, darken the appropriate circle corresponding to the same question number on the OMR sheet.**
4. Darkening two circles for the same question makes the answer invalid.
5. **Damaging/overwriting, using whiteners** on the **OMR** sheets are strictly prohibited.

---

1. Which of the following is a characteristic of hierarchical clustering?
   A) It always requires the number of clusters to be predefined
   B) The dendrogram is used to visualize the hierarchical structure of clusters
   C) Hierarchical clustering is a non – iterative method
   D) Agglomerative hierarchical clustering begins by assuming all data points belong to a single clustering.
   a) (A) & (C) only b) (B) & (D) only
   c) (B), (C) & (D) only d) (A), (B) &(D) only

2. Which of the following statements best describes a Self Organization Map (SOM)?
   a) SOM is a type of supervised learning algorithm used for classification tasks.
   b) SOM reduces the dimensionality of data by mapping high dimensional input to a lower – dimensional grid while preserving topological relationship.
   c) SOM is based on the back propagation algorithm, similar to other neural network
   d) None of these

3. Which of the following statements about dissimilarity measures is correct?
   a) Dissimilarity measures are used to calculate how similar two objects are in clustering problems.
   b) The Euclidean distance is a common dissimilarity measure that calculates the square root of the sum of square difference between corresponding features.
   c) Dissimilarity measures are used to quantify the difference between objects in terms of a similarity metric
   d) None of these

4. What is the primary function of a Self Organizing Map (SOM)?
   a) Classification b) Regression
   c) Clustering and Visualization d) Anomaly detection

5. Which of the following best describes the learning process in a SOM?
   a) Supervised learning with labels
   b) Reinforcement learning
   c) Unsupervised learning using competitive learning
   d) Genetic algorithm based learning

6. In a SOM, which of the following happens during the training process?
   a) Only the Best Matching Unit (BMU) is updated
   b) The weights of all units are updated equally
   c) All neurons receive the same input signal
   d) None of these

7. Which of the following is true about the topological structure in SOMs?
   a) The topological structure is destored during training
   b) Similar data points map to distant neurons
   c) The map preserves the topological relationships between data points
   d) None of these

8. What does a neighborhood function in a SOM control?
   a) The number of neurons in the map
   b) The distance between the data points
   c) The influence of neighboring neurons on the weight update process
   d) The number of layers in the map

9. Which of the following is a common use of data analysis in Civil engineering?
   a) Traffic congestion forecasting b) Structural load analysis
   c) Analyzing tensile strength of materials d) All of these

10. In a structural engineering, which type of data would be most important for conducitng a finite element analysis?
    a) Soil Data
    b) Material properties , geometric dimensions and loading condition
    c) Weather data for construction scheduling
    d) None of these

11. Which of the following best describes data selection in Machine learning?
    a) Choosing relevant features from the data set
    b) Deciding which data samples to include in the training set
    c) Preprocessing the data by removing noise
    d) Deciding on the evaluation metric to use

12. Which of the following is a common technique for selecting a subset of data for model training?
    a) Cross validation b) K – means clustering
    c) Principle component analysis d) Bagging

13. What is "Sampling Bias" in data selection?
    a) The situation where the training data is not representative of the population
    b) When the data is over sampling for specific features
    c) The failure to normalize the data set
    d) A technique used to reduce dimensionality

14. Which method can help in reducing selection bias when selecting data samples for training?
    a) Random sampling b) Standardization
    c) Data augmentation d) Feature scaling

15. Feature selection is the process of :
    a) Identifying and removing irrelevant or redundant features from the data
    b) Scaling features to a common range
    c) Creating new features from the existing ones
    d) Splitting the data into training and test sets

16. In pavement design, the use of the Elastic layered theory primarily focuses on :
    a) Determining the maximum load that the pavement can bear
    b) Predicting to life span of the pavement
    c) Estimating the structural capacity of different pavement layers
    d) Identifying traffic volume patterns

17. Which machine learning technique can be used to predict traffic flow patterns based on historical data in civil engineering?
    a) Decision trees b) Artificial Neural Networks (ANNs)
    c) Linear programming d) Time – series forecasting

18. Which method is typically used in structural health monitoring to detect changes in the behavior of a structure overtime?
    a) Vibration based monitoring b) Cross validation
    c) Image recognition d) particle swarm optimization

19. Which of the following is an example of a machine learning technique that can be used for predictive maintenance in civil engineering?
    a) Decision trees b) Random Forests
    c) Support Vector Machines (SVM) d) All of these

20. Which data analysis method is commonly employed to optimize the design of a bridge in terms of material usage, cost and local capacity?
    a) Multi – objective optimization b) Time series analysis
    c) Random sampling d) Hierarchical clustering

21. Which of the following methods is used to handle missing values in the data set during attribute understanding?
    a) Featuring scale b) Imputation
    c) One hot coding d) Correlation analysis

22. Which type of data transformation is typically required for numerical attributes in machine learning models?
    a) One – hot encoding b) Normalization
    c) Label encoding d) Feature extraction

23. Which of the following is an example of a categorical attribute?
    a) Age b) Income c) Gender d) Temperature

24. One – dimensional (1D) data is typically represented as :
    a) A matrix with multiple rows and columns
    b) A single column or row of values
    c) A cube with depth, width and height
    d) A set of attributes with values linked by relationship

25. One – dimensional (1D) data represents :
    a) A sequence of values from multiple variables
    b) A sequence of values from single variables
    c) Multiple rows and columns of data
    d) A multi dimensional array

26. Which of the following is a measure of the shape of a distribution?
    a) Mean b) Skewness c) Standard deviation d) Variance

27. A data set with a positive skewness indicates that
    a) Most of the data are concentrated on the right side
    b) Most of the data are concentrated on the left side
    c) The data is symmetrical
    d) The data is bimodal

28. Which of the following would likely be true for a perfectly symmetrical distribution?
    a) The mean less than the median
    b) The skewness is 0
    c) The data is positively skewed
    d) The mode is greater than the mean

29. Which of the following is a method to standardize multi dimensional data, especially when the variables are having different scales?
    a) Normalization b) K – mean clustering
    c) Decision tress d) Variance analysis

30. Which of the following is a common method for measuring the spread or variability in multidimensional data?
    a) Covariance Matrix b) Mean Absolute Error (MAE)
    c) Standard error of the mean d) Mode

31. Which of the following frameworks is commonly used in data mining and knowledge?
    a) SEMMA b) CRISP - DM c) Both (a) & (b) d) None of these

32. Which of the following best describes the relationship between knowledge and data?
    a) Data is raw facts, while knowledge is derived from interpreting data
    b) Knowledge and data are the same and can be used interchangeably
    c) Data is processed information , whereas knowledge is unstructured data.
    d) Knowledge is collected first and details generated from it.

33. Which of the following correctly represents the hierarchy from raw data to knowledge?
    a) Knowledge → Information → Data b) Data → Knowledge → Information
    c) Data → Information → Knowledge d) Information → Data → Knowledge

34. Which of the following is NOT a measure of central tendency in descriptive statistics?
a) Mean b) Median c) Mode d) Standard deviation

35. Which of the following is NOT a common criterion for assessing knowledge?
a) Accuracy b) Relevance c) Randomness d) Consistency

36. Which of the following is a key criterion for evaluating the quality of knowledge?
a) Validity b) Ambiguity c) Inconsistency d) Irrelevance

37. Which of the following is the main purpose of inferential statistics?
a) To summarize and describe data
b) To make predications or generalizations about a population based on a sample
c) To organize raw data into labels and graphs
d) To calculate the exact values for an entire population.

38. What is the first step in the data analysis process?
a) Data cleaning b) Data collection
c) Data visualization d) Data interpretation

39. What does the 'S' in SEMMA stand for?
a) Sample b) Structure c) Select d) Survey

40. What does CRISP – DM stand for?
a) Cross – Industry standard process for Data Mining
b) Common Research Information for Statical Models
c) Critical Reporting and Information for Statical Models
d) Custom Research and Information for Data Management

41. Which of the following is an example of a model class?
a) Linear Regression b) Support vector machine
c) Decision tree d) All of these

42. Which of the following is an example of a black – box model?
a) Logistic regression b) Decision tree
c) Random forest d) Linear regression

43. Fitting criteria are used to :
a) Choose the best training data set
b) Evaluate the quality of the model based on the test data
c) Optimize the parameters of the model to fit the data well
d) Select the features for the model

44. Which of the following is commonly used as a fitting criterion for regression problems?
a) Mean Squared Error (MSE) b) Gini Index
c) Entropy d) R - Squared

45. Which of the following is a common score function used for classification problems?
a) Accuracy b) Log - loss c) AUC - ROC d) All of these

46. Which of the following best describes a "Score function" in machine learning?
a) A function that evaluates the performance of a model
b) A function that predicts the target variable
c) A function used to optimize the model parameters
d) A function used to preprocess the data

47. Which of the following is a measure of interestingness in association rule mining?
a) Support b) Mean Squared Error (MSE)
c) R – Squared d) Normalized Mutual Information (NMI)

48. Which of the following is NOT a common measure of interestingness for association rules?
a) Confidence b) Lift c) R - Squared d) Support

49. What is the key advantage of using a closed form solution for model fitting?
a) It guarantees and optimal solution in less time
b) It avoids the need for an iterative process
c) It is suitable for very large data sets
d) It works well for highly non linear data

50. Which of the following is a closed form solution for fitting a linear model?
a) Stochastic Gradient Descent (SGD) b) Newton's method
c) The Normal equation d) Support vector machines

*****

CV 6th Sem.

CBCS SCHEME - Make-Up Exam

BCV657C

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

Question Paper Version : A

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Data Analytics for Civil Engineers

Time: 1 hr Max. Marks: 50

INSTRUCTIONS TO THE CANDIDATES

1. Answer all the **fifty** questions, each question carries one mark.
2. Use only **Black ball point pen** for writing / darkening the circles.
3. **For each question, after selecting your answer, darken the appropriate circle corresponding to the same question number on the OMR sheet.**
4. Darkening two circles for the same question makes the answer invalid.
5. **Damaging/overwriting, using whiteners** on the **OMR** sheets are strictly prohibited.

---

1. A key difference between classification and regression in data analysis is :
   a) Classification predicts categorical outcomes ; regression predicts continuous outcomes
   b) Classification is supervised ; regression in unsupervised
   c) Classification requires large datasets ; regression does not
   d) Classification is part of descriptive statistics ; regression is not

2. Which of the following is an unsupervised data mining task?
   a) Classification b) Regression c) Clustering d) Prediction

3. Which tool is MOST commonly used for exploratory data analysis?
   a) Photoshop b) Excel/Python(pandas, matplotlib, seaborn
   c) AutoCAD d) SQL Server only

4. Which of the following is an example of data analysis task?
   a) Data visualization b) Data collection only
   c) Hardware configuration d) Operating system installation

5. In CRISP-DM, which is the first phase of the process?
   a) Data Preparation b) Business understanding
   c) Modeling d) Evaluation

6. CRISP-DM stand for :
   a) Cross-Industry Standard Processes for Data Mining
   b) Comprehensive Research in Statistical Processing – Data Management
   c) Critical Review of Intelligent Systems for Predictive Data Models
   d) Cross-Industry Standard Planning for Data Management

7. Which of the following correctly represents the steps of the SEMMA process model?
   a) Sample → Explore → Modify → Model →Assess
   b) Search → Estimate → Measure →Model → Apply
   c) Select → Examine → Manage → Model → Assess
   d) Sample → Extract → Merge →Model → Analyze

8. The term Knowledge Discovery in Databases (KDD) refers to :
   a) Database management and optimization
   b) Process of finding valid, and useful patterns in data
   c) Manual extraction of facts from data warehouses
   d) Only data cleaning and preparation

9. Which of the following is a key goal of Exploratory Data Analysis (EDA)?
   a) To prove hypotheses with certainty
   b) To identify patterns, anomalies, or relationship in data visually and statistically
   c) To eliminate the need for statistical modeling
   d) To finalize machine learning models

10. Inferential statistics is primarily concerned with :
    a) Collecting raw data
    b) Describing observed values only
    c) Making generalizations or predictions about a population from sample data
    d) Organizing data into tables and charts

11. Which attribute type allows for true zero ratio comparison?
    a) Nominal b) Ordinal c) Interval d) Ratio

12. Which of the following is an interval attribute?
    a) Age of a bridge
    b) Temperature in Celsius
    c) River discharge rate
    d) Strength of concrete (MPa)

13. The mean, median and mode are example of :
    a) Dispersion measures
    b) Location (central tendency) measures
    c) Shape measures
    d) Normalization techniques

14. Standard deviation is a measures of :
    a) Central tendency b) Dispersion c) Skewness d) Shape

15. Skewness in one-dimensional data indicates :
    a) Central tendency
    b) Spread of values
    c) Asymmetry in distribution
    d) Correlation strength

16. Which measure describes the peakedness or flatness of a distribution?
    a) Mean b) Variance c) Kurtosis d) Range

17. For multidimensional data, which of the following is used to measure similarity?
    a) Mean absolute deviation
    b) Euclidean distance
    c) Variance
    d) Median

18. A density plot is primarily used to :
    a) Show correlation between two variable
    b) Show probability distribution of a variable
    c) Display quartiles and outliers
    d) Summarize multidimensional data

19. A scatter plot is best suited for :
    a) Displaying a single variable's distribution
    b) Analyzing relationship between two variables
    c) Showing skewness and kurtosis
    d) Identifying median and quartiles

20. If correlation between compressive strength and curling time of concrete is 0.92, it indicates :
    a) Weak negative correlation
    b) Strong positive correlation
    c) No correlation
    d) Non-linear relation only

21. The four general of modeling are :
    a) Selection → Building → Training → Deployment
    b) Selection → Fitting → Validation → Deployment
    c) Collection → Preprocessing → Deployment → Validation
    d) Training → Testing → Cross-validation → Reporting

22. Model classes in data modeling are typically categorized into :
    a) Statistical, machine learning, and physical models
    b) Block-box and white-box models
    c) Deterministic and probabilistic models
    d) All of these

23. A commonly used fitting criterion in regression models is :
    a) Mean Squared Error (MSE)
    b) Lift ratio
    c) Entropy
    d) Gini index

24. In classification problems, the error function most commonly used is :
    a) Cross-entropy loss
    b) R-squared value
    c) Root mean square error
    d) Adjusted R^2

25. The measure of interestingness in association rule mining is often:
    a) Accuracy and Precision
    b) Support and Confidence
    c) Variance and Kurtosis
    d) Mean and Median

26. An example of a closed-form algorithm in regression is :
    a) Gradient descent method
    b) Normal equation for linear regression
    c) Back propagation in neural networks
    d) K-means clustering

27. Type I error in classification refers to :
a) Accepting a false hypothesis (false positive)
b) Rejecting a false hypothesis (true negative)
c) Accepting true hypothesis (true positive)
d) Rejecting a true hypothesis (false negative)

28. In civil engineering, predictive models are often used to :
a) Estimate traffic flow and pavement deterioration
b) Generate soil types randomly
c) Replace structural analysis entirely
d) Avoid statistical sampling

29. A block-box model used in civil engineering for predicting rainfall-runoff relationship is
a) Linear regression
b) ANN (Artificial Neural Networks)
c) PCA
d) Box plot

30. In flood risk modeling, Type I error may correspond to :
a) Predicting flood when there is no flood (false alarm)
b) Predicting no flood when there is flood
c) Correctly predicting flood event
d) Ignoring rainfall data

31. Feature selection in data preparation aims to :
a) Remove irrelevant and redundant attributes
b) Create artificial data points
c) Increase dimensionality
d) Normalize the entire dataset

32. Selecting a top-ranked subset of data generally done using:
a) Feature importance ranking
b) Random sampling
c) Dummy coding
d) PCA decomposition only

33. The cross-product method in feature selection refers to :
a) Generating all possible combinations of features
b) Multiplying attributes values
c) Splitting dataset into equal parts
d) Correlation calculation only

34. The wrapper approach in features selection evaluates features by :
a) Ranking them using correlation only
b) Testing subsets of features against a predictive model
c) Removing features randomly
d) Using PCA to reduce data

35. A correlation-based filter for features selection keeps attributes that :
a) Have low correlation with target variable
b) Have high correlation with target and low correlation with each other
c) Are random but evenly distributed
d) Have highest mean values

36. The process of data cleaning focuses on :
a) Removing irrelevant features
b) Correcting errors, inconsistencies and duplicates
c) Increasing dataset size
d) Creating synthesis features

37. Which of the following is NOT a common method to handle missing values?
a) Deletion of records
b) Mean/Median imputation
c) Regression imputation
d) Randomization of missing values

38. Ensuring operability in data preparation means :
a) Data can be used effectively in downstream models
b) Data is stored in external hardware only
c) Features are ranked in descending order
d) All missing values are ignored

39. Assuring impartiality in data preparation refers to :
a) Removing bias data collection and preprocessing
b) Making dataset more skewed
c) Ignoring categorical attributes
d) Using only numerical data

40. In traffic engineering, data cleaning is most important for :
a) Removing duplicate vehicle counts or sensor errors
b) Randomly increasing vehicle counts
c) Reducing lane widths in data models
d) Ignoring traffic flow trends

41. Clustering in data mining is mainly used for :
a) Predicting future values
b) Grouping similar data points without prior labels
c) Testing statistical hypotheses
d) Determining regression coefficients

42. In agglomerative hierarchical clustering, the process starts with :
a) One big cluster and splits into smaller cluster
b) Each data point as its own cluster merges step by step
c) Random partitioning of data
d) A fixed number of clusters predefined

43. Which distance metric is also known as the L1 norm?
a) Euclidean
b) Manhattan
c) Chebyshev
d) Cosine

44. The Chebyshev distance between two points is defined as :
a) Maximum absolute difference along any dimension
b) Sum of squared differences
c) Square root of squared differences
d) Angle between vectors

45. Consine similarity is most appropriate when data is represented as :
a) Categorical variable
b) Magnitude-based values
c) High-dimensional sparse vectors
d) Low-dimensional coordinates only

46. Deviation measures in clustering are used to :
a) Find the centroid of clusters
b) Measure how far objects deviate from patterns
c) Merge clusters in hierarchical methods
d) Normalize attribute values

47. In association rule mining, the rule X → Y means :
a) If Y occurs, X always occurs
b) If X occurs, Y tends to occur
c) X and Y are independent
d) X excludes Y

48. Self-Organizing Maps (SOMs) are a type of :
a) Supervised neural network
b) Unsupervised neural network
c) Regression model
d) Clustering index

49. The main goal of a Self-Organizing Map is to :
a) Classify categorical variables
b) Reduce high-dimensional data into low-dimensional grids
c) Estimate regression coefficients
d) Generate random clusters

50. Association rule mining could be applied in transportation engineering to :
a) Find relationships between traffic flow patterns and accident frequency
b) Randomly split traffic data
c) Generate noise in GPS records
d) Eliminate seasonal variations

* * * * *

CBCS SCHEME

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

BCV613B

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Design of Formwork and Scaffolding

Time: 3 hrs. Max. Marks: 100

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. M : Marks , L: Bloom's level , C: Course outcomes.

		Module – 1	M	L	C
Q.1	a.	Write a note on selection of formwork and formwork materials.	10	L2	CO1
	b.	List the various types of formwork materials used in the construction.	10	L1	CO1
		OR			
Q.2	a.	Explain briefly on assembling and de-shuttering of formwork system.	10	L2	CO1
	b.	Explain eh following : i) Formwork for columns ii) Formwork for beams.	10	L2	CO1
		Module – 2			
Q.3	a.	List the various design assumptions and design methods considered in formwork.	10	L2	CO2
	b.	Explain the various steps followed involved in slab formwork.	10	L2	CO2
		OR			
Q.4	a.	List out the various design assumptions and design methods considered in the formwork.	10	L2	CO2
	b.	List the various steps involved in slab formwork and checks.	10	L2	CO2
		Module – 3			
Q.5	a.	List and explain the objectives of formwork schedule.	10	L2	CO3
	b.	What is the role of mobilization distribution in formwork. Explain in brief.	10	L1	CO3
		OR			
Q.6	a.	Explain the purpose of BOQ in formwork.	10	L2	CO3
	b.	List the factors involved in mobilization distribution of formwork.	10	L2	CO3
		Module – 4			
Q.7	a.	Enumerate advantages and limitations of Special formwork.	10	L2	CO4
	b.	List the various steps involved in formwork of high rise construction.	10	L2	CO4
		OR			
Q.8	a.	Explain the various steps involved in shuttering and de-shuttering special formwork.	10	L2	CO4
	b.	Explain the following : i) Ladding safety ii) Loading classification in special formwork.	10	L2	CO4
		Module – 5			
Q.9	a.	Explain various types failures in formwork.	10	L2	CO1
	b.	List the various safety principals considered in formwork.	10	L2	CO1
		OR			
Q.10	a.	Explain the process of formwork erection and safety.	10	L2	CO1
	b.	Briefly explain the prevention in formwork failures.	10	L2	CO1

* * * * *

CBCS SCHEME

BCV657C

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

Question Paper Version : A

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Data Analytics for Civil Engineers

Time: 1 hr.] [Max. Marks: 50

INSTRUCTIONS TO THE CANDIDATES

1. Answer all the **fifty** questions, each question carries one mark.
2. Use only **Black ball point pen** for writing / darkening the circles.
3. **For each question, after selecting your answer, darken the appropriate circle corresponding to the same question number on the OMR sheet.**
4. Darkening two circles for the same question makes the answer invalid.
5. **Damaging/overwriting, using whiteners** on the **OMR** sheets are strictly prohibited.

---

1. Which of the following frameworks is commonly used in data mining and knowledge?
   a) SEMMA b) CRISP - DM c) Both (a) & (b) d) None of these

2. Which of the following best describes the relationship between knowledge and data?
   a) Data is raw facts, while knowledge is derived from interpreting data
   b) Knowledge and data are the same and can be used interchangeably
   c) Data is processed information , whereas knowledge is unstructured data.
   d) Knowledge is collected first and details generated from it.

3. Which of the following correctly represents the hierarchy from raw data to knowledge?
   a) Knowledge → Information → Data b) Data → Knowledge → Information
   c) Data → Information → Knowledge d) Information → Data → Knowledge

4. Which of the following is NOT a measure of central tendency in descriptive statistics?
   a) Mean b) Median c) Mode d) Standard deviation

5. Which of the following is NOT a common criterion for assessing knowledge?
   a) Accuracy b) Relevance c) Randomness d) Consistency

6. Which of the following is a key criterion for evaluating the quality of knowledge?
   a) Validity b) Ambiguity c) Inconsistency d) Irrelevance

7. Which of the following is the main purpose of inferential statistics?
   a) To summarize and describe data
   b) To make predications or generalizations about a population based on a sample
   c) To organize raw data into labels and graphs
   d) To calculate the exact values for an entire population.

8. What is the first step in the data analysis process?
   a) Data cleaning b) Data collection
   c) Data visualization d) Data interpretation

9. What does the 'S' in SEMMA stand for?
   a) Sample b) Structure c) Select d) Survey

10. What does CRISP – DM stand for?
    a) Cross – Industry standard process for Data Mining
    b) Common Research Information for Statical Models
    c) Critical Reporting and Information for Statical Models
    d) Custom Research and Information for Data Management

11. Which of the following methods is used to handle missing values in the data set during attribute understanding?
    a) Featuring scale b) Imputation
    c) One hot coding d) Correlation analysis

12. Which type of data transformation is typically required for numerical attributes in machine learning models?
    a) One – hot encoding b) Normalization
    c) Label encoding d) Feature extraction

13. Which of the following is an example of a categorical attribute?
    a) Age b) Income c) Gender d) Temperature

14. One – dimensional (1D) data is typically represented as :
    a) A matrix with multiple rows and columns
    b) A single column or row of values
    c) A cube with depth, width and height
    d) A set of attributes with values linked by relationship

15. One – dimensional (1D) data represents :
    a) A sequence of values from multiple variables
    b) A sequence of values from single variables
    c) Multiple rows and columns of data
    d) A multi dimensional array

16. Which of the following is a measure of the shape of a distribution?
    a) Mean b) Skewness c) Standard deviation d) Variance

17. A data set with a positive skewness indicates that
    a) Most of the data are concentrated on the right side
    b) Most of the data are concentrated on the left side
    c) The data is symmetrical
    d) The data is bimodal

18. Which of the following would likely be true for a perfectly symmetrical distribution?
    a) The mean less than the median
    b) The skewness is 0
    c) The data is positively skewed
    d) The mode is greater than the mean

19. Which of the following is a method to standardize multi dimensional data, especially when the variables are having different scales?
a) Normalization b) K – mean clustering
c) Decision tress d) Variance analysis

20. Which of the following is a common method for measuring the spread or variability in multidimensional data?
a) Covariance Matrix b) Mean Absolute Error (MAE)
c) Standard error of the mean d) Mode

21. Which of the following is an example of a model class?
a) Linear Regression b) Support vector machine
c) Decision tree d) All of these

22. Which of the following is an example of a black – box model?
a) Logistic regression b) Decision tree
c) Random forest d) Linear regression

23. Fitting criteria are used to :
a) Choose the best training data set
b) Evaluate the quality of the model based on the test data
c) Optimize the parameters of the model to fit the data well
d) Select the features for the model

24. Which of the following is commonly used as a fitting criterion for regression problems?
a) Mean Squared Error (MSE) b) Gini Index
c) Entropy d) R - Squared

25. Which of the following is a common score function used for classification problems?
a) Accuracy b) Log - loss c) AUC - ROC d) All of these

26. Which of the following best describes a "Score function" in machine learning?
a) A function that evaluates the performance of a model"
b) A function that predicts the target variable
c) A function used to optimize the model parameters
d) A function used to preprocess the data

27. Which of the following is a measure of interestingness in association rule mining?
a) Support b) Mean Squared Error (MSE)
c) R – Squared d) Normalized Mutual Information (NMI)

28. Which of the following is NOT a common measure of interestingness for association rules?
a) Confidence b) Lift c) R - Squared d) Support

29. What is the key advantage of using a closed form solution for model fitting?
a) It guarantees and optimal solution in less time
b) It avoids the need for an iterative process
c) It is suitable for very large data sets
d) It works well for highly non linear data

30. Which of the following is a closed form solution for fitting a linear model?
a) Stochastic Gradient Descent (SGD) b) Newton's method
c) The Normal equation d) Support vector machines

31. Which of the following best describes data selection in Machine learning?
a) Choosing relevant features from the data set
b) Deciding which data samples to include in the training set
c) Preprocessing the data by removing noise
d) Deciding on the evaluation metric to use

32. Which of the following is a common technique for selecting a subset of data for model training?
a) Cross validation b) K – means clustering
c) Principle component analysis d) Bagging

33. What is "Sampling Bias" in data selection?
a) The situation where the training data is not representative of the population
b) When the data is over sampling for specific features
c) The failure to normalize the data set
d) A technique used to reduce dimensionality

34. Which method can help in reducing selection bias when selecting data samples for training?
a) Random sampling b) Standardization
c) Data augmentation d) Feature scaling

35. Feature selection is the process of :
a) Identifying and removing irrelevant or redundant features from the data
b) Scaling features to a common range
c) Creating new features from the existing ones
d) Splitting the data into training and test sets

36. In pavement design, the use of the Elastic layered theory primarily focuses on :
a) Determining the maximum load that the pavement can bear
b) Predicting to life span of the pavement
c) Estimating the structural capacity of different pavement layers
d) Identifying traffic volume patterns

37. Which machine learning technique can be used to predict traffic flow patterns based on historical data in civil engineering?
a) Decision trees b) Artificial Neural Networks (ANNs)
c) Linear programming d) Time – series forecasting

38. Which method is typically used in structural health monitoring to detect changes in the behavior of a structure overtime?
a) Vibration based monitoring b) Cross validation
c) Image recognition d) particle swarm optimization

39. Which of the following is an example of a machine learning technique that can be used for predictive maintenance in civil engineering?
    a) Decision trees
    b) Random Forests
    c) Support Vector Machines (SVM)
    d) All of these

40. Which data analysis method is commonly employed to optimize the design of a bridge in terms of material usage, cost and local capacity?
    a) Multi – objective optimization
    b) Time series analysis
    c) Random sampling
    d) Hierarchical clustering

41. Which of the following is a characteristic of hierarchical clustering?
    A) It always requires the number of clusters to be predefined
    B) The dendrogram is used to visualize the hierarchical structure of clusters
    C) Hierarchical clustering is a non – iterative method
    D) Agglomerative hierarchical clustering begins by assuming all data points belong to a single clustering.
    a) (A) & (C) only
    b) (B) & (D) only
    c) (B), (C) & (D) only
    d) (A), (B) &(D) only

42. Which of the following statements best describes a Self Organization Map (SOM)?
    a) SOM is a type of supervised learning algorithm used for classification tasks.
    b) SOM reduces the dimensionality of data by mapping high dimensional input to a lower – dimensional grid while preserving topological relationship.
    c) SOM is based on the back propagation algorithm, similar to other neural network
    d) None of these

43. Which of the following statements about dissimilarity measures is correct?
    a) Dissimilarity measures are used to calculate how similar two objects are in clustering problems.
    b) The Euclidean distance is a common dissimilarity measure that calculates the square root of the sum of square difference between corresponding features.
    c) Dissimilarity measures are used to quantify the difference between objects in terms of a similarity metric
    d) None of these

44. What is the primary function of a Self Organizing Map (SOM)?
    a) Classification
    b) Regression
    c) Clustering and Visualization
    d) Anomaly detection

45. Which of the following best describes the learning process in a SOM?
    a) Supervised learning with labels
    b) Reinforcement learning
    c) Unsupervised learning using competitive learning
    d) Genetic algorithm based learning

46. In a SOM, which of the following happens during the training process.
    a) Only the Best Matching Unit (BMU) is updated
    b) The weights of all units are updated equally
    c) All neurons receive the same input signal
    d) None of these

47. Which of the following is true about the topological structure in SOMs?
    a) The topological structure is destored during training
    b) Similar data points map to distant neurons
    c) The map preserves the topological relationships between data points
    d) None of these

48. What does a neighborhood function in a SOM control?
    a) The number of neurons in the map
    b) The distance between the data points
    c) The influence of neighboring neurons on the weight update process
    d) The number of layers in the map

49. Which of the following is a common use of data analysis in Civil engineering?
    a) Traffic congestion forecasting
    b) Structural load analysis
    c) Analyzing tensile strength of materials
    d) All of these

50. In a structural engineering, which type of data would be most important for conducitng a finite element analysis?
    a) Soil Data
    b) Material properties , geometric dimensions and loading condition
    c) Weather data for construction scheduling
    d) None of these

* * * * *

CBCS SCHEME - Make-Up Exam

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

BCV654C

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Integrated Waste Management for a Smart City

Time: 3 hrs. Max. Marks: 100

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. M : Marks , L: Bloom's level , C: Course outcomes.

		Module – 1	M	L	C
Q.1	a.	Explain the characteristics of municipal solid wastes.	10	L2	CO1
	b.	Discuss the factors that affect the generation of solid wastes.	10	L2	CO1
		OR			
Q.2	a.	Explain Integrated solid waste management. Explain its issues.	10	L2	CO1
	b.	With sketch, explain stationary container systems.	10	L2	CO1
		Module – 2			
Q.3	a.	Enlist factors affecting the composting process and explain any two in detail.	10	L2	CO1
	b.	Explain about Bangalore method of composting and draw a neat sketch.	10	L2	CO1
		OR			
Q.4	a.	Outline a flow chart showing the steps involved in the aerobic compositing process. Explain the factors affecting composting process.	10	L2	CO2
	b.	Explain Swachh Bharat Mission and Smart City Program.	10	L2	CO2
		Module – 3			
Q.5	a.	What are the types of gasifier,s? Formulate the advantages of biogasification.	10	L2	CO2
	b.	Categorize the current issues in solid waste management.	10	L2	CO2
		OR			
Q.6	a.	What are the types of energy recovery process and explain any one process.	10	L2	CO2
	b.	Illustrate the current status of MSW management in each of the 20 smart cities.	10	L2	CO3
		Module – 4			
Q.7	a.	Explain the strategies for effective C and D waste management.	10	L2	CO3
	b.	Categorize the opportunities and solutions of using C and D waste.	10	L2	CO3
		OR			
Q.8	a.	List and explain the steps involved in beneficial Reuse of Construction and demolition waste.	10	L2	CO3
	b.	Interpret the purpose of construction and demolition waste management.	10	L2	CO3

BCV654C

		Module – 5			
Q.9	a.	What is E-waste? What are the impacts of E-waste on the environment?	10	L2	CO4
	b.	Explain the importance E-waste management.	10	L2	CO4
		OR			
Q.10	a.	What is the current status of e-waste in India?	10	L2	CO4
	b.	Explain the E-waste management rules 2016.	10	L2	CO4

*****