18340

BRAINWARE UNIVERSITY

Term End Examination 2025-2026
Programme – M.Tech.(CSE)-AIML-2025
Course Name – Machine Learning Applications
Course Code - MTA10402
(Semester I)

Full Marks : 60 **Time : 2:30 Hours**

[The figure in the margin indicates full marks. Candidates are required to give their answers in their own words as far as practicable.]

Group-A
(Multiple Choice Type Question) 1 x 15=15

1. *Choose the correct alternative from the following :*

(i) Identify the purpose of bagging in ensemble learning

a) Reduce bias
b) Reduce variance by training on bootstrapped datasets
c) Increase training data
d) Use single deep tree

(ii) Implement a probability-based decision for binary classification

a) Use standard deviation
b) Calculate posterior probability
c) Find loss function
d) Apply grid search

(iii) Implement structured kernel for graph-based classification

a) Ignore edge weights
b) Use common node labels
c) Count graph isomorphisms
d) Compare substructure patterns

(iv) Apply latent semantic indexing in search engine optimization

a) Increase ad coverage
b) Compute similarity using semantic space
c) Use click prediction model
d) Tokenize content

(v) Examine the benefit of kernel trick in high-dimension space

a) Slows down optimization
b) Needs full transformation
c) Avoids explicit mapping and saves computation
d) Always overfits

(vi) Summarize the role of latent variables in PLSA

a) Represent observed features
b) Introduce noise
c) Explain co-occurrence
d) Increase sparsity

(vii) Explain why graphical models reduce computational complexity

a) Increase nodes
b) Use nested loops
c) Encode independencies
d) Skip likelihood terms

(viii) Analyze how latent variables improve text modeling in PLSA

a) Increase feature overlap
b) Flatten distribution

c) Capture topic semantics | d) Decrease accuracy

(ix) Evaluate trade-offs in using Gaussian processes for large datasets

a) Fast to train | b) High storage and compute cost
c) Ignore covariance | d) Apply gradient descent only

(x) Apply filter methods to select top features from financial dataset

a) Use PCA components | b) Apply Chi-square test
c) Use SGD classifier | d) Normalize data

(xi) Apply wrapper method to optimize feature set

a) Use regular expression | b) Use SVM-RFE
c) Filter nulls | d) Drop all columns

(xii) Assess ethics of biased feature selection in security algorithms

a) Makes system neutral | b) Enhances fairness
c) Introduces discrimination | d) Reduces complexity

(xiii) Apply LDA to distinguish two classes of biological samples

a) Merge all data into one class | b) Calculate inter-class variance
c) Use K-means | d) Encode all values as 0

(xiv) Identify ICA application for biomedical signal processing

a) Increase training time | b) Reconstruct EKG noise
c) Extract independent EEG signals | d) Segment video frames

(xv) Create a dimensionality reduction module for text summarization

a) Count all nouns | b) Skip pre-processing
c) Extract feature vectors using LSA | d) Remove all verbs

Group-B

(Short Answer Type Questions) 3 x 5=15

2. Define supervised learning with a suitable example. (3)
3. Analyse application of Kernel PCA for reducing the dimensionality of a non-linear dataset. (3)
4. Illustrate how Bayes' theorem updates model belief after new evidence. (3)
5. Use CCA to identify correlations between two sets of variables. (3)
6. Create a system that selects between LDA and PCA based on the data label availability. (3)

OR

Devise a hybrid approach combining CCA and ICA for multimodal data analysis. (3)

Group-C

(Long Answer Type Questions) 5 x 6=30

7. Apply the principles of computational learning theory to assess generalization ability of ML models. (5)
8. Implement a kernel-based SVM model for classifying a medical dataset and explain each step. (5)
9. Evaluate the role of kernel PCA in face recognition tasks. (5)
10. Examine the conversion of a Bayesian network to a Markov network through moralization, and discuss effects on inference algorithms. (5)
11. Construct a multi-stage feature extraction process using LDA and NMF for image analysis. (5)
12. Evaluate feature selection methods in scientific research data preprocessing. (5)

OR

Compare embedded, filter, and wrapper methods for high-dimensional genomic datasets. (5)

****