



MARATHWADA MITRA MANDAL'S

COLLEGE OF ENGINEERING

An Autonomous Institute | Approved by AICTE New Delhi
Recognised by Directorate of Technical Education Mumbai
Affiliated to Savitribai Phule Pune University

- NAAC Accredited with A++ Grade
- NBA Accredited
- 'Best College Award 2019' by SPPU

Department of Artificial Intelligence and Data Science

Class: B.E.

Subject: Computer Laboratory I (Part A: Machine Learning)

Assignment No. 1: Feature Transformation

1. What is Principal Component Analysis, and what problem does it solve?
2. How does PCA differ from feature selection?
3. What is the role of eigenvalues and eigenvectors in PCA?
4. Why is it important to standardize features before applying PCA?
5. How is the number of principal components decided in PCA and in LDA?
6. Can PCA be used for classification tasks? Why or why not?
7. What is Linear Discriminant Analysis and how is it different from PCA? In what scenarios would you prefer LDA over PCA?
8. How does LDA achieve class separability?
9. What are the assumptions behind LDA?

Assignment No. 2: Regression Analysis

1. What is the difference between Linear, Ridge, and Lasso Regression?
2. How does regularization help prevent overfitting?
3. What is the role of the regularization parameter λ ?
4. Why does Lasso perform feature selection while Ridge does not?
5. How do you interpret R^2 and RMSE in model evaluation?
6. What is multicollinearity, and how do Ridge and Lasso address it?
7. Why is it important to standardize features before applying Ridge or Lasso?
8. When would you prefer Lasso over Ridge or Linear Regression?

Assignment No. 3: Classification Analysis

1. What is the difference between classification and regression in supervised learning?
2. Define confusion matrix with components and Classification report.
3. What do accuracy, precision, recall, and F1-score represent in classification tasks?
4. How does increasing or decreasing the value of K affect the performance of the KNN algorithm?
5. Explain the bias-variance tradeoff in the context of classification models.
6. What are the common challenges in classification problems with imbalanced datasets?
7. How does the KNN algorithm classify a new data point? What distance metric is typically used?
8. Why is feature scaling important before applying KNN?





MARATHWADA MITRA MANDAL'S

COLLEGE OF ENGINEERING

An Autonomous Institute | Approved by AICTE New Delhi
Recognised by Directorate of Technical Education Mumbai
Affiliated to Savitribai Phule Pune University

- NAAC Accredited with A++ Grade
- NBA Accredited
- 'Best College Award 2019' by SPPU

Assignment No. 4: Clustering Analysis

1. What is clustering in machine learning, and how does it differ from classification?
2. What are the main assumptions and objectives of the K-Means algorithm?
3. Explain the concept of centroids in K-Means. How are they updated during training?
4. What is the elbow method and how is it used to determine the optimal number of clusters?
5. What is the silhouette score and how does it help in evaluating clustering performance?
6. What are the limitations of K-Means clustering? How does initialization affect the result?
7. Can K-Means handle non-spherical clusters or outliers effectively? Why or why not?
8. How does K-Medoids differ from K-Means in terms of centroid selection?

Assignment No. 5: Ensemble Learning

1. What is ensemble learning and why is it used in machine learning?
2. Differentiate between bagging and boosting techniques.
3. What are the advantages of using ensemble methods over individual models?
4. How does combining weak learners result in a strong learner in ensemble methods?
5. How does a Random Forest differ from a single Decision Tree?
6. Explain the role of bootstrapping and feature randomness in Random Forest.
7. What is Out-of-Bag (OOB) error in Random Forest and how is it useful?

Assignment No. 6: Reinforcement Learning (RL)

1. What is reinforcement learning and how does it differ from supervised and unsupervised learning?
2. Define the following terms in the context of reinforcement learning: agent, environment, state, action, reward.
3. What is a policy in RL? How is it different from a value function?
4. Explain the exploration vs exploitation trade-off in RL. How is this typically managed?
5. What is the Bellman equation and what role does it play in RL?
6. How is the environment defined in the Tic-Tac-Toe or Taxi problem in terms of state and action space?
7. How do terminal states affect the training of an RL agent in a game like Tic-Tac-Toe?

