DESCRIPTIVE :

1 a. Discuss the importance of feature selection in Machine Learning. Compare filter, wrapper, and embedded methods for feature selection.

1. b. Describe Principal Components Analysis (PCA) as a dimension reduction technique. Explain its working, the role of eigenvalues and eigenvectors, and its implications for model performance and interpretability.

2. a. Explain decision trees in Machine Learning for regression and classification. Discuss their pros and cons compared to linear models.

b. Evaluate the suitability of decision trees and ensemble methods (e.g., Random Forests and Boosting) for predicting customer churn in a subscription-based service. Describe how ensemble methods combine multiple decision trees and their impact on model performance, interpretability, and generalization

3. a. Define unsupervised learning and its role in Machine Learning. Explain the concept of clustering and provide a brief overview of its main applications.

b. Apply SVM to the identification of fraudulent credit card transactions. Cover kernel functions and the trade-offs between the margin and misclassification error

1.a. Explain the role of tuning parameters in Machine Learning. Provide an example.

b. Discuss the benefits of Principal Components Analysis (PCA) for dimension reduction, especially in high-dimensional data situations

2.a.Define decision trees and their types and advantages.

b.Describe the construction of Random Forests and discuss an application scenario.

3.a.Define Support Vector Machines (SVM) and support vectors.

b.Describe practical considerations in clustering using methods like K-Means and Hierarchical Clustering with a Usecase.

1.a.Explain the role of tuning parameters in Machine Learning. Provide an example.

b.Discuss the benefits of Principal Components Analysis (PCA) for dimension reduction, especially in high-dimensional data situations.

2.a.Define decision trees and their types and advantages.

b.Describe the construction of Random Forests and discuss an application scenario.

3.a.Define unsupervised learning and the maximal margin concept in SVM. Provide an example of SVM's application with non-linear decision boundaries.

b.Describe practical considerations in clustering using methods like K-Means and Hierarchical Clustering with a Usecase.

OPEN BOOK :

1. Under what circumstances do you opt for decision tree models in Machine Learning? Discuss their advantages and disadvantages compared to other modeling techniques, such as linear regression.

2. Discuss the significance and advantages of Bayesian Additive Regression Trees (BART) in Machine Learning. How does BART differ from other regression techniques, and in which situations is it particularly effective? Provide an example of a scenario where BART was applied to yield valuable results

3. Provide an in-depth explanation of how Support Vector Machines (SVM) work, emphasizing the concept of a maximal margin classifier. Discuss the advantages of SVM in handling non-linear decision boundaries. Give a practical example where SVMs proved effective for classification in your research or the industry.

1. What are the key considerations when selecting features for a Machine Learning model? How does feature selection contribute to model accuracy and interpretability? Provide a specific case from your research or the field.

2. When is it appropriate to use Random Forests as a Machine Learning technique, and what types of data are most suitable for Random Forest models? Share real-world use cases that demonstrate the benefits of Random Forests.

3. Define unsupervised learning and describe various models within this category. Explain the role of unsupervised learning in data analysis and pattern discovery.