**Savitribai Phule Pune University**

**Fourth Year of Computer Engineering (2019 Course)**

**410246: Laboratory Practice III**

**Examination Scheme:**

**Term work: 50 Marks**

**Practical: 50 Marks**

**Companion Course:**

**Machine Learning(410242),**

**Design and Analysis of Algorithms (410241),**

**Blockchain Technology(410243)**

**MOCK PRACTICAL ON 9th/11th/12th/13th Oct 2023 :**

**PICK UP RANDOMLY ONE PS OUT OF FOLLOWING 12 ML PS**

**IN FIRST 1 HOUR IMPLEMENT SUCCESSFULLY ELSE REPEAT**

**LP III : Machine Learning(410242)**

**PS1 : Predict the price of the Uber ride from a given pickup point to the agreed drop-off location.**

**Perform following tasks:**

**1. Pre-process the dataset.**

**2. Identify outliers.**

**3. Check the correlation.**

**4. Implement linear regression and random forest regression models.**

**5. Evaluate the models and compare their respective scores like R2, RMSE, etc.**

**Dataset link: https://www.kaggle.com/datasets/yasserh/uber-fares-dataset**

**PS2 : Predict the house price**

**Perform following tasks:**

**1. Pre-process the dataset.**

**2. Identify outliers.**

**3. Check the correlation.**

**4. Implement linear regression**

**5. Evaluate the models and compare their respective scores like R2, RMSE, etc.**

**Dataset link:**

[**https://www.kaggle.com/datasets/iamsouravbanerjee/house-rent-prediction-dataset**](https://www.kaggle.com/datasets/iamsouravbanerjee/house-rent-prediction-dataset)

**PS3 :**

**Classify the email using the binary classification method. Email Spam detection has two states:**

**a) Normal State – Not Spam**

**b) Abnormal State – Spam.**

**Use K-Nearest Neighbors for classification.**

**Analyze their performance.**

**Dataset link:**

**https://www.kaggle.com/datasets/balaka18/email-spam-classification-dataset-csv**

**PS4 : Read titanic dataset, predict for following test cases whether person will survive or not. Assume necessary features if necessary.**

**1) Class : 1, Gender : Female, Age : 42, Siblings / Spouses : 2, Parents/Children: 1**

**2) Class : 3, Gender : male, Age : 30, Siblings / Spouses : 0, Parents/Children: 0**

**Dataset link: https://www.kaggle.com/c/titanic/data**

**PS5: Implement Gradient Descent Algorithm to find the local minima of a function. For example, find the**

**local minima of the function y=(x+3)2 starting from the point x=2.**

**PS6:**

**Implement K-Nearest Neighbors algorithm on diabetes.csv dataset. Compute confusion matrix, accuracy, error rate, precision and recall on the given dataset.**

**Dataset link : https://www.kaggle.com/datasets/abdallamahgoub/diabetes**

**PS7:**

**Implement Naive Bayes algorithm on diabetes.csv dataset. Compute confusion matrix, accuracy, error rate, precision and recall on the given dataset.**

**Dataset link : https://www.kaggle.com/datasets/abdallamahgoub/diabetes**

**PS8:**

**Implement Decision Tree algorithm on diabetes.csv dataset. Compute confusion matrix, accuracy, error rate, precision and recall on the given dataset.**

**Dataset link :** [**https://www.kaggle.com/datasets/abdallamahgoub/diabetes**](https://www.kaggle.com/datasets/abdallamahgoub/diabetes)

**PS9:**

**Implement Random Forest algorithm on diabetes.csv dataset. Compute confusion matrix, accuracy, error rate, precision and recall on the given dataset.**

**Dataset link :** [**https://www.kaggle.com/datasets/abdallamahgoub/diabetes**](https://www.kaggle.com/datasets/abdallamahgoub/diabetes)

**PS10 :**

**Classify the email using the binary classification method. Email Spam detection has two states:**

**a) Normal State – Not Spam**

**b) Abnormal State – Spam.**

**Use Support Vector Machine for classification.**

**Analyze their performance.**

**Dataset link:**

**https://www.kaggle.com/datasets/balaka18/email-spam-classification-dataset-csv**

**PS11 :**

**Classify the email using the binary classification method. Email Spam detection has two states:**

**a) Normal State – Not Spam**

**b) Abnormal State – Spam.**

**Use Logistic Regression for classification.**

**Analyze their performance.**

**Dataset link:**

**https://www.kaggle.com/datasets/balaka18/email-spam-classification-dataset-csv**

**PS12 :**

**Implement K-Means clustering/ hierarchical clustering on sales\_data\_sample.csv dataset. Determine the number of clusters using the elbow method.**

**Dataset link : https://www.kaggle.com/datasets/kyanyoga/sample-sales-data**

**MOCK PRACTICAL ON 9th/11th/12th/13th Oct 2023 :**

**PICK UP RANDOMLY ONE PS OUT OF FOLLOWING 5 DAA PS**

**IN NEXT 1 HOUR IMPLEMENT SUCCESSFULLY ELSE REPEAT**

**DAA ASSIGNMENTS**

**PS13 : Write a program non-recursive and recursive program to calculate Fibonacci numbers and analyze their time and space complexity.**

**PS14 : Write a program to implement Huffman Encoding using a greedy strategy.**

**PS15 : Write a program to solve a 0-1 Knapsack problem using dynamic programming or branch and**

**bound strategy.**

**PS16 : Design n-Queens matrix having first Queen placed. Use backtracking to place remaining**

**Queens to generate the final n-queen‘s matrix.**

**PS17 : Write a program for analysis of quick sort by using deterministic and randomized variant.**

**MOCK PRACTICAL ON 9th/11th/12th/13th Oct 2023 :**

**PICK UP RANDOMLY ONE PS OUT OF FOLLOWING 2 BCT PS**

**IN LAST 1 HOUR IMPLEMENT SUCCESSFULLY ELSE REPEAT**

**Blockchain Technology**

**PS18 : Write a smart contract on a test network, for Bank account of a customer for following**

**operations:**

** Deposit money**

** Withdraw Money**

** Show balance**

**PS19 : Write a program in solidity to create Student data. Use the following constructs:**

** Structures**

** Arrays**

** Fallback**

**Deploy this as smart contract on Ethereum and Observe the transaction fee and Gas values.**