Set-01

Q1. This data set contains details of a startups and the target variable is a continuous variable. Dataset contains input features and predicting profit. Apply necessary data preprocessing Construct Linear Regression Model using dependent variable (R&D Spend, Marketing Spend, State) to predict profit based on given data using Jupyter Notebook.

Q2. This data set contains details of a disease diagnostic parameter. Construct two clusters using K Means Algorithm to predict whether new patient will have that disease or not using Jupyter notebook and Deploy model on local host using Flask. Use Gender, Glucose, BP, Skin Thickness, Insulin, BMI, Pedigree function, age.

Set-02

Q1.This data set contains details of a startups and the target variable is a continuous variable. Dataset contains input features and predicting profit. Apply necessary data preprocessing Construct Random Forest Model using dependent variable (R&D Spend , Administration , State ) to predict profit based on given data using Jupyter Notebook.

Q2. This data set contains details of a disease diagnostic parameter. Construct two clusters using Hierarchical Clustering Algorithm to predict whether new patient will have that disease or not using Jupyter notebook and Deploy model on local host using Flask. Use Gender, Glucose, BP, Skin Thickness, Insulin, BMI, Pedigree function, age.

Set-03

Q1.This data set contains details of a startups and the target variable is a continuous variable. Dataset contains input features and predicting profit. Apply necessary data preprocessing Construct SVM Model using dependent variable (Marketing Spend , Administration ,State ) to predict profit based on given data using Jupyter Notebook.

Q2. This data set contains details of a disease diagnostic parameter and the target variable is a binary variable reflecting the fact whether the person has that disease or not. Find three most significant Principal components and Construct naïve bayes Classifier using these three pcs to predict whether new patient will have that disease or not using Jupyter notebook and Deploy model on local host using Flask. Use Gender, Glucose, BP, Skin Thickness, Insulin, BMI, Pedigree function, age.

Set-04

Q1.This data set contains details of a startups and the target variable is a continuous variable. Dataset contains input features and predicting profit. Apply necessary data preprocessing Construct Naïve Bayes using dependent variable (R&D Spend, Marketing Spend, Administration, and State) to predict profit based on given data using Jupyter Notebook.

Q2. Find out Frequent item sets and Association rule for 50% support and 50% confidence using Apriori algorithm for given dataset This data set contains list of items for each transaction. Use Jupyter and Deploy model on local host using Flask

Set-05

Q1. This data set contains details of a bank's customers and the target variable is a binary variable reflecting the fact whether the customer left the bank (closed his account) or he continues to be a customer. Apply necessary data preprocessing. Construct Decision Tree Classifier to predict whether new customer will leave the bank or not using Jupyter Notebook. choose following columns as independent variable in dataset.(CreditScore, Geography, Gender, Age, Tenure, HasCrCard, IsActiveMember, EstimatedSalary).

Q2. This database is to identify a voice as male or female, based upon acoustic properties of the voice and speech. Construct two clusters using K Means Algorithm to predict whether the voice is of male or female using Jupyter notebook and Deploy model on local host using Flask

Set-06

Q1. This data set contains details of a bank's customers and the target variable is a binary variable reflecting the fact whether the customer left the bank (closed his account) or he continues to be a customer. Apply necessary data preprocessing. Construct KNN Classifier to predict whether new customer will leave the bank or not using Jupyter Notebook. choose following columns as independent variable in dataset.( CreditScore, Gender, Age,Tenure, Balance,IsActiveMember, EstimatedSalary).

Q2. This database is to identify a voice as male or female, based upon acoustic properties of the voice and speech. Construct two clusters using Hierarchical Clustering to predict whether the voice is of male or female using Jupyter notebook and Deploy model on local host using Flask.

Set-07

Q1. This data set contains details of a bank's customers and the target variable is a binary variable reflecting the fact whether the customer left the bank (closed his account) or he continues to be a customer.Apply necessary data preprocessing. Construct SVM Classifier to predict whether new customer will leave the bank or not using Jupyter Notebook. choose following columns as independent variable in dataset.( CreditScore, Gender, Age, Tenure, Balance, HasCrCard, EstimatedSalary)

Q2. This database is to identify a voice as male or female, based upon acoustic properties of the voice and speech Find three most significant Principal components and Construct Decision tree Classifier using these three PCs to predict whether voice belong to male or female using Jupyter notebook and Deploy model on local host using Flask.

Set-08

Q1. This data set contains details of a bank's customers and the target variable is a binary variable reflecting the fact whether the customer left the bank (closed his account) or he continues to be a customer.Apply necessary data preprocessing. Construct Random Forest to predict whether new customer will leave the bank or not using Jupyter Notebook. choose following columns as independent variable in dataset.( CreditScore, Geography, Age, Tenure, Balance, IsActiveMember, EstimatedSalary)

Q2. Find out Frequent item sets and Association rule for 50% support and 50% confidence using Apriori algorithm for given data set. This Dataset contains list of items for each transaction. Use Jupyter and Deploy model on local host using Flask.

Set-09

Q1. This data set contains details of a heart patients and the target variable is a binary variable reflecting the fact whether the patient has heart disease or not. Apply necessary data preprocessing Construct Decision Tree to predict whether new patient will have heart disease or not using Jupyter Notebook. choose following columns as independent variable in dataset . Use Age, Sex, chestPain, RestingBP, Cholestrol, FastingBS, RestingECG, OldPEak,ST\_slope.

Q2. This data set contains details of a bank's customers to predict whether new customer will leave the bank or not. Construct two clusters using K Means Algorithm to using Jupyter notebook and Deploy model on local host using Flask. Choose following columns as independent variable in dataset.(CreditScore, Geography, Gender, Age, Tenure, HasCrCard, IsActiveMember,EstimatedSalary)

Set-10

Q1.This data set contains details of a heart patients and the target variable is a binary variable reflecting the fact whether the patient has heart disease or not. Apply necessary data preprocessing. Construct Naïve Bayes to predict whether new patient will have heart disease or not using Jupyter Notebook. choose following columns as independent variable in dataset. Use Age, ex, ChestPaintType, Resting BP, Cholestrol, Fasting BS, MaxHR, Exercise Angina, Old peak, St\_Slope.

Q2. This data set contains details of a bank's customers to predict whether new customer will leave the bank or not. Construct two clusters using Hierarchical Clustering Algorithm to using Jupyter notebook and Deploy model on local host using Flask. Choose following columns as independent variable in dataset. (CreditScore, Geography, Gender, Age, Tenure, HasCrCard, IsActiveMember,EstimatedSalary)

Set-11

Q1. This data set contains details of a heart patients and the target variable is a binary variable reflecting the fact whether the patient has heart disease or not.Apply necessary data preprocessing. Construct Random Forest to predict whether new patient will have heart disease or not using Jupyter Notebook. choose following columns as independent variable in dataset. Use Age, Sex, chest Pain Type, Resting BP, cholesterol, RestingECG, MAxHR, ExerciseAngina, OldPeak ,Stslope.

Q2. This data set contains details of a bank's customers to predict whether new customer will leave the bank or not. Find three most significant Principal components and Construct SVM Classifier using these three pcs to predict whether new patient will have that disease or not using Jupyter notebook and Deploy model on local host using Flask. Choose following columns as independent variable in dataset. (CreditScore, Geography, Gender, Age, Tenure, HasCrCard, IsActiveMember,EstimatedSalary)

Set-12

Q1. This data set contains details of a heart patients and the target variable is a binary variable reflecting the fact whether the patient has heart disease or not. Apply necessary data preprocessing. Construct KNN to predict whether new patient will have heart disease or not using Jupyter Notebook. choose following columns as independent variable in dataset. Use Age, Sex, chest Pain Type, Resting BP, cholesterol, FastingBS, Resting ECG, Exercise Angina,STSlope.

Q2. Find out Frequent item sets and Association rule for 60% support and 60% confidence using Apriori algorithm for given data set. This Dataset contains list of items for each transaction. Use Jupyter and Deploy model on local host using Flask.

Set-13

Q1. This data set contains details of a disease diagnostic parameter and the target variable is a binary variable reflecting the fact whether the person has that disease or not. Apply necessary data preprocessing. Construct Decision Tree Classifier to predict whether new patient will have that disease or not using Jupyter Notebook. choose following columns as independent variable in dataset (Gender, Glucose, BP ,Skin Thickness, Insulin, BMI ,Pedigree Function ,Age)

Q2. This data set contains details of a heart patients and the target variable is a binary variable reflecting the fact whether the patient has heart disease or not. Construct two clusters using K Means Algorithm to using Jupyter notebook and Deploy model on local host using Flask

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Age | Sex | ChestPainType | RestingBP | Cholesterol | FastingBS | MaxHR | ExerciseAngina | Oldpeak | ST\_Slope |

Set-14

Q1. This data set contains details of a disease diagnostic parameter and the target variable is a binary variable reflecting the fact whether the person has that disease or not.Apply necessary data preprocessing. Construct KNN to predict whether new patient will have that disease or not using Jupyter Notebook. choose following columns as independent variable in dataset(Glucose, BP, Skin Thickness, Insulin, BMI, Age)

Q2. This data set contains details of a heart patients and the target variable is a binary variable reflecting the fact whether the patient has heart disease or not. Construct two clusters using Hierarchical Clustering Algorithm to using Jupyter notebook and Deploy model on local host using Flask Choose following columns as independent variable in dataset.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Age | Sex | ChestPainType | RestingBP | Cholesterol | FastingBS | MaxHR | ExerciseAngina | Oldpeak | ST\_Slope |

Set-15

Q1. This data set contains details of a disease diagnostic parameter and the target variable is a binary variable reflecting the fact whether the person has that disease or not. Apply necessary data preprocessing. Construct SVM to predict whether new patient will have that disease or not using Jupyter Notebook. choose following columns as independent variable in dataset. ( Pedigree Function, Insulin, Skin Thickness, BP, Age)

Q2. This data set contains details of a heart patients and the target variable is a binary variable reflecting the fact whether the patient has heart disease or not. Find four most significant Principal components and Construct SVM Classifier using these four pcs to predict whether new patient will have that heart disease or not using Jupyter notebook and Deploy model on local host using Flask Choose following columns as independent variable in dataset. **(10+15+5 Marks)**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Age | Sex | ChestPainType | RestingBP | Cholesterol | FastingBS | MaxHR | ExerciseAngina | Oldpeak | ST\_Slope |

Set-16

Q1. This data set contains details of a disease diagnostic parameter and the target variable is a binary variable reflecting the fact whether the person has that disease or not.Apply necessary data preprocessing. Construct Random Forest to predict whether new patient will have that disease or not using Jupyter Notebook. choose following columns as independent variable in dataset.( BMI, Insulin, Glucose, Gender, BP, Age)

Q2. Find out Frequent item sets and Association rule for 50% support and 50% confidence using Apriori algorithm for given data set. This Dataset contains list of items for each transaction. Use Jupyter and Deploy model on local host using Flask.

Set-17

Q1. This data set contains details of a disease diagnostic parameter and the target variable is a binary variable reflecting the fact whether the person has that disease or not. Apply necessary data preprocessing. Construct Naïve Bayes Classifier to predict whether new patient will have that disease or not using Jupyter Notebook. choose following columns as independent variable in dataset.(Gender, Glucose,BP,Skin Thickness,BMI,Pedigree Function, Age)

Q2. This data set contains details of a houses and the target variable is a continuous variable. Dataset contains input features and predicting house prices. Construct two clusters using K Means Algorithm to using Jupyter notebook and Deploy model on local host using Flask. Choose following columns as independent variable in dataset.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| area | bathrooms | stories | mainroad | guestroom | hotwaterheating | parking | prefarea | furnishingstatus |

Set-18

Q1. This data set contains details of a disease diagnostic parameter and the target variable is a binary variable reflecting the fact whether the person has that disease or not. Apply necessary data preprocessing . Construct Decision Tree to predict whether new patient will have that disease or not using Jupyter Notebook. choose following columns as independent variable in dataset. (Gender, Glucose, BP, Skin Thickness, Insulin, BMI, Age)

Q2. This data set contains details of a houses and the target variable is a continuous variable. Dataset contains input features and predicting house prices. Construct two clusters Hierarchical Clustering Algorithm to using Jupyter notebook and Deploy model on local host using Flask. Choose following columns as independent variable in dataset.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| area | bathrooms | stories | mainroad | guestroom | hotwaterheating | parking | prefarea | furnishingstatus |

Set-19

Q1. This data set contains details of a disease diagnostic parameter and the target variable is a binary variable reflecting the fact whether the person has that disease or not. Apply necessary data preprocessing .Construct Random Forest to predict whether new patient will have that disease or not using Jupyter Notebook. choose following columns as independent variable in dataset (Gender, Glucose, BP, Skin Thickness, Insulin, BMI, Age)

Q2. This data set contains details of a houses and the target variable is a continuous variable. Dataset contains input features and predicting house prices. Find three most significant Principal components and Construct Linear Regression using these three pcs to predict prices using Jupyter notebook and Deploy model on local host using Flask. Choose following columns as independent variable in dataset.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| area | bathrooms | stories | mainroad | guestroom | hotwaterheating | parking | prefarea | furnishingstatus |

Set-20

Q1. This data set contains details of a disease diagnostic parameter and the target variable is a binary variable reflecting the fact whether the person has that disease or not.Apply necessary data preprocessing. Construct SVM to predict whether new patient will have that disease or not using Jupyter Notebook. choose following columns as independent variable in dataset (Gender, Glucose, BP, Skin Thickness, Insulin, BMI, Age)

Q2. Find out Frequent item sets and Association rule for 50% support and 50% confidence using Apriori algorithm for given data set. This Dataset contains list of items for each transaction. Use Jupyter and Deploy model on local host using Flask

Set-21

Q1.This data set contains details of a houses and the target variable is a continuous variable. Dataset contains input features and predicting house prices. Apply necessary data preprocessing -Construct Linear Regression Model using dependent variable to predict house prices based on given data using Jupyter Notebook. Use Area, bathrooms, stories, main road, guestroom, hotwater heating, parking, prefarea, furnishing status.

Q2. This data set contains details of a startups and the target variable is a continuous variable. Dataset contains input features and predicting profit. Construct two clusters using K Means Algorithm using independent variable (R&D Spend , Marketing Spend , State ) to predict profit based on given data using Jupyter notebook and Deploy model on local host using Flask

Set-22

Q1.This data set contains details of a houses and the target variable is a continuous variable. Dataset contains input features and predicting house prices.Apply necessary data preprocessing -Construct SVM Model using dependent variable to predict house prices based on given data using Jupyter Notebook. Use area, bedrooms, Bathrooms. Stories, main road, guestroom, Air conditioning, parking, Furnishing status.

Q2. This data set contains details of a startups and the target variable is a continuous variable. Dataset contains input features and predicting profit. Construct two clusters using Hierarchical Clustering Algorithm using independent variable (R&D Spend , Marketing Spend , State ) to predict profit based on given data using Jupyter notebook and Deploy model on local host using Flask

Set-23

Q1.This data set contains details of a houses and the target variable is a continuous variable. Dataset contains input features and predicting house prices. Apply necessary data preprocessing -Construct KNN Model using dependent variable to predict house prices based on given data using Jupyter Notebook. Use area, bedrooms ,mainroad. guestroom, Hot water heating, Air conditioning, prefarea, Furnishing status.

Q2. This data set contains details of a startups and the target variable is a continuous variable. Dataset contains input features and predicting profit. Find three most significant Principal components and Construct Linear Regression using these three pcs to predict profit using Jupyter notebook and Deploy model on local host using Flask.

Set-24

Q1. This data set contains details of a houses and the target variable is a continuous variable. Dataset contains input features and predicting house prices. Apply necessary data preprocessing -Construct Naïve Bayes Model using dependent variable to predict house prices based on given data using Jupyter Notebook. Use area, bedrooms, Bathrooms. Stories, Basement, Hot water heating, Air conditioning, parking, Furnishing status.

Q2. Find out Frequent item sets and Association rule for 50% support and 50% confidence using Apriori algorithm for given data set. This Dataset contains list of items for each transaction. Use Jupyter and Deploy model on local host using Flask.