

Department of Computer Science

MSc (CS) -II Sem-III Machine Learning Assignments

Assignments	Start Date	End Date	Sign
Assignment 1 <ol style="list-style-type: none">1. Write a python program to Prepare Scatter Plot (Use Forge Dataset / Iris Dataset).2. Write a python program to find all null values from a given data set and remove them.3. Write a python program the Categorical values in numeric format for a given dataset.4. Take iris flower dataset and reduce 4D data to 2D data using PCA. Then train the model and predict new flowers with given measurements.			
Assignment 2 <ol style="list-style-type: none">1. Write a python program to implement multiple Linear Regression for predicting house price. Divide dataset into train and test data while giving it to model and predict prices of house.2. Write a python program to implement logistic Regression. Use dataset crash.csv is an accident survivor's dataset portal for USA hosted by data.gov. The dataset contains passengers age and speed of vehicle (mph) at the time of impact and fate of passengers (1 for survived and 0 for not survived) after a crash. Use logistic regression to decide if the age and speed can predict the survivability of the passengers.3. Write a python program to fit the simple linear regression and polynomial linear regression models to Salary_positions.csv data. Find which one is more accurately fitting to the given data. Also predict the salaries of level 11 and level 12 employees.			
Assignment 3 <ol style="list-style-type: none">1. Write a python program to Implement Decision Tree classifier model onData which is extracted from images that were taken from genuine and forged banknote-like specimens. (refer UCI dataset https://archive.ics.uci.edu/dataset/267/banknote+authentication)2. Classify the iris flowers dataset using SVM and find out the flower type depending on the given input data like sepal length, sepal width, petal length and petal width. Find accuracy of all SVM kernels.3. Write a python program to implement k-nearest Neighbors ML algorithm to build prediction model. (Iris Dataset / Salary dataset)4. Write a python program to Implement Naïve Bayes algorithm.			

5. Implement Non-linear regression model (Decision Tree, SVM, KNN) to predict the consumption of petrol use petrol consumption dataset. (https://www.kaggle.com/code/ajinkyaa/linear-regression-petrol-consumption)			
<p style="text-align: center;">Assignment 4</p> <ol style="list-style-type: none"> 1. Use K-means clustering model and classify the employees into various income groups or clusters. Preprocess data if require (i.e. drop missing or null values). Use elbow method and Silhouette Score to find value of k. 2. Write a python program to implement Agglomerative clustering on a synthetic dataset. 3. The data set refers to clients of a wholesale distributor. It includes the annual spending in monetary units on diverse product categories. Using data Wholesale customer dataset compute agglomerative clustering to find out annual spending clients in the same region. https://archive.ics.uci.edu/dataset/292/wholesale+customers 			
<p>Reference datasets:</p> <p>https://www.kaggle.com/datasets/manmohan291/housepricedata.csv</p> <p>https://www.kaggle.com/uciml/iris</p> <p>https://www.kaggle.com/code/prasadperera/the-boston-housing-dataset/input</p> <p>https://www.kaggle.com/datasets/irfanarullah/groceries</p> <p>https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database</p> <p>https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset</p> <p>https://www.kaggle.com/datasets/gurdit559/canada-per-capita-income-single-variable-data-set</p> <p>https://www.kaggle.com/datasets/vaibhavxnn/google-stock-prices-training-and-test-data</p> <p>https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents</p> <p>https://archive.ics.uci.edu/datasets</p>			

Department of Computer Science

MSc (CA) -I Sem-II Machine Learning Assignments

Assignments	Start Date	End Date	Sign
Assignment 1 5. Write a python program to Prepare Scatter Plot (Use Forge Dataset / Iris Dataset). 6. Write a python program to find all null values from a given data set and remove them. 7. Write a python program the Categorical values in numeric format for a given dataset. 8. Take iris flower dataset and reduce 4D data to 2D data using PCA. Then train the model and predict new flower with given measurements.			
Assignment 2 4. Write a python program to implement multiple Linear Regression for predicting house price. Divide dataset into train and test data while giving it to model and predict prices of house. 5. Write a python program to implement logistic Regression. Use dataset crash.csv is an accident survivor's dataset portal for USA hosted by data.gov. The dataset contains passengers age and speed of vehicle (mph) at the time of impact and fate of passengers (1 for survived and 0 for not survived) after a crash. Use logistic regression to decide if the age and speed can predict the survivability of the passengers. 6. Write a python program to fit the simple linear regression and polynomial linear regression models to Salary_positions.csv data. Find which one is more accurately fitting to the given data. Also predict the salaries of level 11 and level 12 employees.			
Assignment 3 6. Write a python program to Implement Decision Tree classifier model onData which is extracted from images that were taken from genuine and forged banknote-like specimens. (refer UCI dataset https://archive.ics.uci.edu/dataset/267/banknote+authentication) 7. Classify the iris flowers dataset using SVM and find out the flower type depending on the given input data like sepal length, sepal width, petal length and petal width. Find accuracy of all SVM kernels. 8. Write a python program to implement k-nearest Neighbors ML algorithm to build prediction model. (Iris Dataset / Salary dataset) 9. Write a python program to Implement Naïve Bayes.			

10. Implement Non-linear regression model (Decision Tree, SVM, KNN) to predict the consumption of petrol use petrol consumption dataset. (https://www.kaggle.com/code/ajinkyaa/linear-regression-petrol-consumption)			
<p style="text-align: center;">Assignment 4</p> <p>4. Use K-means clustering model and classify the employees into various income groups or clusters. Preprocess data if require (i.e. drop missing or null values). Use elbow method and Silhouette Score to find value of k.</p> <p>5. Write a python program to implement Agglomerative clustering on a synthetic dataset.</p> <p>6. The data set refers to clients of a wholesale distributor. It includes the annual spending in monetary units on diverse product categories. Using data Wholesale customer dataset compute agglomerative clustering to find out annual spending clients in the same region. https://archive.ics.uci.edu/dataset/292/wholesale+customers</p>			
<p>Reference datasets:</p> <p>https://www.kaggle.com/datasets/manmohan291/housepricedata.csv</p> <p>https://www.kaggle.com/uciml/iris</p> <p>https://www.kaggle.com/code/prasadperera/the-boston-housing-dataset/input</p> <p>https://www.kaggle.com/datasets/irfanarullah/groceries</p> <p>https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database</p> <p>https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset</p> <p>https://www.kaggle.com/datasets/gurdit559/canada-per-capita-income-single-variable-data-set</p> <p>https://www.kaggle.com/datasets/vaibhavxnn/google-stock-prices-training-and-test-data</p> <p>https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents</p> <p>https://archive.ics.uci.edu/datasets</p>			

Department of Computer Science

MSc (DS) -I Sem-II Machine Learning Assignments

Assignment	Start Date	End Date	Sign
Assignment 1 1. Write a python program to Prepare Scatter Plot (Use Forge Dataset / Iris Dataset). 2. Write a python program to find all null values from a given data set and remove them. 3. Write a python program the Categorical values in numeric format for a given dataset. 4. Write a python program to transform data with Principal Component Analysis (PCA).	22/1/24	27/1/24	
Assignment 2 5. Write a python program to implement simple Linear Regression for predicting house price. 6. Write a python program to implement multiple Linear Regression for a given dataset. 7. Write a python program to implement logistic Regression for a given dataset. 8. Write a python program to implement Polynomial Regression for given dataset.	29/1/24	3/2/24	

<p style="text-align: center;">Assignment 3</p> <p>9. Write a python program to Implement Decision Tree Model.</p> <p>10. Write a python program to implement linear SVM for Regression and Classification.</p> <p>11. Write a python program to implement k-nearest Neighbors ML algorithm to build prediction model. (Iris Dataset / Salary dataset)</p> <p>12. Write a python program to Implement Naïve Bayes.</p>	12/2/24	20/2/24	
<p style="text-align: center;">Assignment 4</p> <p>13. Write a python program to implement k-means algorithm on a synthetic dataset.</p> <p>14. Write a python program to implement Agglomerative clustering on a synthetic dataset.</p>	21/2/24	29/3/24	

Assignment 1

1. Write a python program to Prepare Scatter Plot (Use Forge Dataset / Iris Dataset).
 2. Write a python program to find all null values from a given data set and remove them.
 3. Write a python program the Categorical values in numeric format for a given dataset.
- Write a python program to transform data with Principal Component Analysis (PCA). Consider handwritten digit dataset.

Write a python program to implement simple Linear Regression for predicting house price.

Write a python program to implement multiple Linear Regression for predicting house price.

Write a python program to implement logistic Regression for predicting whether a person will buy the insurance_data.csv. Use insurance_data.csv

Write a python program to implement logistic Regression for handwritten digit dataset.

Write a python program to implement Polynomial Regression for given dataset. Use position_sal.csv.

Write a python program to Implement Decision Tree Model for classification. Use Decision_Tree_Dataset.csv

Write a python program to implement linear SVM for Regression. Use position_sal.csv

Write a python program to implement linear SVM for Classification. Use iris.csv

Write a python program to implement k-nearest Neighbors algorithm to build prediction model. Use Iris Dataset.

Write a python program to Implement Naïve Bayes for classification. Use titanic.csv/spam.csv dataset

Write a python program to implement k-means algorithm on a synthetic dataset.

Write a python program to implement Agglomerative clustering on a synthetic dataset.