Term Project 03

Contents | Page No.

Term Project 3

Topics:

1. Logistic Regression and Adaptive Boosting Classification
2. Naïve Bayes Classification and Random Forest Classification
3. Decision Tree Regression and Adaptive Boosting Regression
4. K Nearest Neighbor Classification and Decision Tree Classification
5. Linear Regression and Random Forest Regression

Implementation Instructions:

1. Scikit-Learn Python library can be used
2. Comparison between two approaches has to be done
3. Visualization should be done for comparison of approaches
4. Data should be scaled if deemed necessary
5. K-fold cross validation must be used
6. At least four performance metric score have to be used

Instructions for reporting:

1. Strictly formatted
2. Clearly defined input
3. Major steps of processing
4. Accurate output

Description of Datasets

- **BostonHousing.csv**

  Associated Task: Regression
  Number of Instances: 506
  Number of attributes: 14
  Attribute Information:
  1. CRIM - per capita crime rate by town
  2. ZN - proportion of residential land zoned for lots over 25,000 sq.ft.
  3. INDUS - proportion of non-retail business acres per town.
  4. CHAS - Charles River dummy variable (1 if tract bounds river; 0 otherwise)
  5. NOX - nitric oxides concentration (parts per 10 million)
  6. RM - average number of rooms per dwelling
  7. AGE - proportion of owner-occupied units built prior to 1940
  8. DIS - weighted distances to five Boston employment centres
  9. RAD - index of accessibility to radial highways
  10. TAX - full-value property-tax rate per $10,000
  11. PTRATIO - pupil-teacher ratio by town
  12. B - 1000(Bk - 0.63)^2 where Bk is the proportion of blacks by town
  13. LSTAT - % lower status of the population
  14. MEDV - Median value of owner-occupied homes in $1000's

- **HeartDisease.csv**

  Associated Task: Classification
  Number of Instances: 303
  Number of attributes: 14
  Attribute Information:
  1. age: age in years
  2. sex: (1 = male; 0 = female)
  3. cp: chest pain type
  4. trestbps: resting blood pressure (in mm Hg on admission to the hospital)
  5. chol: serum cholestoral in mg/dl
  6. fbs: (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)
  7. restecg: resting electrocardiographic results
  8. thalach: maximum heart rate achieved
  9. exang: exercise induced angina (1 = yes; 0 = no)
  10. oldpeak: ST depression induced by exercise relative to rest
  11. slope: the slope of the peak exercise ST segment
  12. ca: number of major vessels (0-3) colored by flourosopy
  13. tha: l3 = normal; 6 = fixed defect; 7 = reversable defect
  14. target: 1 or 0

- **RedWineQuality.csv**

  Associated Tasks: Classification, Regression

Characteristics: Multivariate
Number of Instances: 1599
Number of attributes: 12
Attribute Information:

Input variables

1. fixed acidity
2. volatile acidity
3. citric acid
4. residual sugar
5. chlorides
6. free sulfur dioxide
7. total sulfur dioxide
8. density
9. pH
10. sulphates
11. alcohol

Output variable

12. quality (score between 0 and 10)

- **LiverPatient.csv**

Associated Task: Classification
Number of Instances: 583
Number of attributes: 11
Attribute Information:

1. age: Age of the patient in years
2. gender: Patient Gender: Male or Female
3. TB: Total Bilirubin
4. DB: Direct Bilirubin
5. alkphos: Alkaline Phosphotase
6. sgpt: Alamine Aminotransferase
7. sgot: Aspartate Aminotransferase
8. TP: Total Proteins
9. ALB: Albumin
10. A_G: Ratio of Albumin and Globulin
11. class: Predictor Class: 1 if patient has Liver Disease and 2 if they do not

- **WhiteWineQuality.csv**

Associated Tasks: Classification, Regression
Characteristics: Multivariate
Number of Instances: 4898
Number of attributes: 12
Attribute Information:

Input variables

1. fixed acidity
2. volatile acidity

3. citric acid
4. residual sugar
5. chlorides
6. free sulfur dioxide
7. total sulfur dioxide
8. density
9. pH
10. sulphates
11. alcohol

Output variable

12. quality (score between 0 and 10)

Course No: CSE 4108
Course Title: Artificial Intelligence Lab
Lab Group: A1

Term Project 03 Topics Allotment

Sl.	Student ID	Allotted Topic
01	160104001	Topic 1
02	160104002	Topic 2
03	160104004	Topic 3
04	160104005	Topic 4
05	160104006	Topic 5
06	160104007	Topic 1
07	160104008	Topic 2
08	160104009	Topic 3
09	160104010	Topic 4
10	160104011	Topic 5
11	160104012	Topic 1
12	160104014	Topic 2
13	160104016	Topic 3
14	160104017	Topic 4
15	160104018	Topic 5
16	160104019	Topic 1
17	160104020	Topic 2
18	160104021	Topic 3
19	160104022	Topic 4
20	160104023	Topic 5
21	160104024	Topic 1
22	160104025	Topic 2
23	160104027	Topic 3

Course No: CSE 4108
Course Title: Artificial Intelligence Lab
Lab Group: A2

Term Project 03 Topics Allotment

Sl.	Student ID	Allotted Topic
01	160104028	Topic 1
02	160104030	Topic 2
03	160104031	Topic 3
04	160104032	Topic 4
05	160104033	Topic 5
06	160104034	Topic 1
07	160104035	Topic 2
08	160104036	Topic 3
09	160104037	Topic 4
10	160104038	Topic 5
11	160104039	Topic 1
12	160104040	Topic 2
13	160104041	Topic 3
14	160104042	Topic 4
15	160104043	Topic 5
16	160104044	Topic 1
17	160104045	Topic 2
18	160104046	Topic 3
19	160104047	Topic 4
20	160104048	Topic 5
21	160104049	Topic 1
22	160104050	Topic 2
23	160104052	Topic 3
24	160104053	Topic 4
25	160104054	Topic 5

Course No: CSE 4108
Course Title: Artificial Intelligence Lab
Lab Group: B1

Term Project 03 Topics Allotment

Sl.	Student ID	Allotted Topic
01	160104055	Topic 1
02	160104056	Topic 2
03	160104057	Topic 3
04	160104058	Topic 4
05	160104059	Topic 5
06	160104060	Topic 1
07	160104061	Topic 2
08	160104063	Topic 3
09	160104064	Topic 4
10	160104065	Topic 5
11	160104066	Topic 1
12	160104067	Topic 2
13	160104068	Topic 3
14	160104070	Topic 4
15	160104072	Topic 5
16	160104073	Topic 1
17	160104074	Topic 2
18	160104076	Topic 3
19	160104077	Topic 4
20	160104078	Topic 5
21	160104079	Topic 1
22	160104080	Topic 2
23	160104081	Topic 3

Course No: CSE 4108
Course Title: Artificial Intelligence Lab
Lab Group: B2

Term Project 03 Topics Allotment

Sl.	Student ID	Allotted Topic
01	160104083	Topic 1
02	160104084	Topic 2
03	160104085	Topic 3
04	160104086	Topic 4
05	160104087	Topic 5
06	160104088	Topic 1
07	160104089	Topic 2
08	160104090	Topic 3
09	160104091	Topic 4
10	160104092	Topic 5
11	160104093	Topic 1
12	160104094	Topic 2
13	160104095	Topic 3
14	160104096	Topic 4
15	160104097	Topic 5
16	160104098	Topic 1
17	160104099	Topic 2
18	160104100	Topic 3
19	160104101	Topic 4
20	160104102	Topic 5
21	160104103	Topic 1
22	160104106	Topic 2
23	160104107	Topic 3
24	160104108	Topic 4

Course No: CSE 4108
Course Title: Artificial Intelligence Lab
Lab Group: C1

Term Project 03 Topics Allotment

Sl.	Student ID	Allotted Topic
01	130204049	Topic 1
02	150204018	Topic 2
03	160104109	Topic 3
04	160104110	Topic 4
05	160104111	Topic 5
06	160104112	Topic 1
07	160104114	Topic 2
08	160104115	Topic 3
09	160104116	Topic 4
10	160104117	Topic 5
11	160104118	Topic 1
12	160104119	Topic 2
13	160104120	Topic 3
14	160104124	Topic 4
15	160104125	Topic 5
16	160104126	Topic 1
17	160104127	Topic 2
18	160104131	Topic 3
19	160104132	Topic 4
20	160104133	Topic 5
21	160104134	Topic 1
22	160104135	Topic 2

Course No: CSE 4108
Course Title: Artificial Intelligence Lab
Lab Group: C2

Term Project 03 Topics Allotment

Sl.	Student ID	Allotted Topic
01	130104134	Topic 1
02	140104074	Topic 2
03	140104141	Topic 3
04	150104110	Topic 4
05	150204010	Topic 5
06	150204027	Topic 1
07	160104136	Topic 2
08	160104138	Topic 3
09	160104139	Topic 4
10	160104140	Topic 5
11	160104141	Topic 1
12	160104142	Topic 2
13	160104143	Topic 3
14	160104144	Topic 4
15	160104145	Topic 5
16	160104146	Topic 1
17	160104147	Topic 2
18	160104148	Topic 3
19	160104149	Topic 4
20	160104150	Topic 5