Subject: 19CSE305

Lab Session: 06

Notes:

1. Please read the assignment notes carefully and comply with the guidelines provided.
2. Code should be checked into GitHub and the report to TurnItIn. Once done, please submit your assignments in Teams.
3. Code non-availability in GitHub shall be marked as zero.
4. Any content copy (statements, figures, codes etc.) from anywhere shall attract a penalty of 10 marks. If you obtain content from anywhere for illustration purposes, please cite the source to avoid penalty.
5. Snapshot / screenshot of code and results not allowed in the report. You may copy content from your own code & results and add it to the report.
6. Provide data, code snippets or illustrations to support your answer, as applicable.

Please use the data associated with your own project.

Refer:

- **https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html**
- **https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html**
- **https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingRegressor.html**
- **https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.ElasticNet.html**
- **https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDRegressor.html#sklearn.linear_model.SGDRegressor**
- **https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html**
- **https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.BayesianRidge.html**
- **https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.ARDRegression.html**
- **https://scikit-learn.org/stable/modules/generated/sklearn.kernel_ridge.KernelRidge.html**
- **https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lasso.html**

Main Section (Mandatory):

A1. Take any 2 features from your project having has real numeric values. Make a scatter plot of the data and observe the pattern.

A2. Create a linear regression model on this data. Consider using one feature as independent variable while the other as dependent variable (you may also round this number to integer). After the model is created, calculate the mean square error by predicting the values from the model. Refer site: https://scikit-learn.org/stable/auto_examples/linear_model/plot_ols.html.

A3. Study all the parameters and attributes associated with linear regression model.

A4. Using the training set available for your project, train a logistic regression classifier. Use this classifier to evaluate your test set accuracy. Study the various parameters associated with logistic regression model and the role they play in the model training.

[NOTE]: For Logistic regression, use bi-class classification problem. If your data has multiple classes, take any two classes. If your data has value prediction problem (Ex: stock price or car value prediction), convert it into a bi-class classification problem and fit a logistic regressor.

A5. Use a Regression Tree and k-NN Regressor for value prediction. Use the data employed for A1.

Optional Section:

O1. Study the above provided regressor links and understand the regression techniques (Gradient Boost, XG Boost, ElasticNet, SGD, SVM, Bayesian Ridge, ARD, Kernel Ridge, LASSO etc.). Make a comparative analysis of their performance on your dataset.

Report Assignment:

1. Search to identify and download more relevant papers for your project. Study them and update the Literature survey section as appropriate.
2. Using the experiments conducted in this lab exercise and results obtained, update the methodology, results & discussions and conclusion portions of your report.