ECS734 Machine Learning for Visual Data Analysis
Lab 3: Age Estimation by Regression

Name : Tanmaiyii Rao
Student ID : 140361229

- This coursework is on age estimation from facial images.
- The directories for the lab files are created in which there is data_age.mat and lab3.m – the file that is to be executed.

Steps

1. The data_age.mat file is loaded.
2. The built-in Matlab function regress() is called that takes the training features and labels as input and learns a linear regression model.
3. The test data is read and the learned linear regression model is applied to estimate the age for each test data point.
4. The MAE and CS value (with a cumulative error level of 5) is computed by comparing the estimated ages with the ground truth ages. The following code has been implemented:

```
w_lr = regress(ytrain,xtrain);
```

```
%% Testing
xtest = teData.feat; % feature
ytest = teData.label; % labels

yhat_test = xtest * w_lr;
```

```
%% Compute the MAE and CS value (with cumulative error level of 5) for linear regression
error_lr = abs(yhat_test-ytest);
mae_lr = sum(error_lr)/size(ytest, 1);
i =5;
CS = sum(error_lr < i == 1)/ size(ytest, 1);

fprintf('MAE(Linear regression(OLS) = %f\n', mae_lr);
fprintf('CS(5) = %f\n', CS);
```

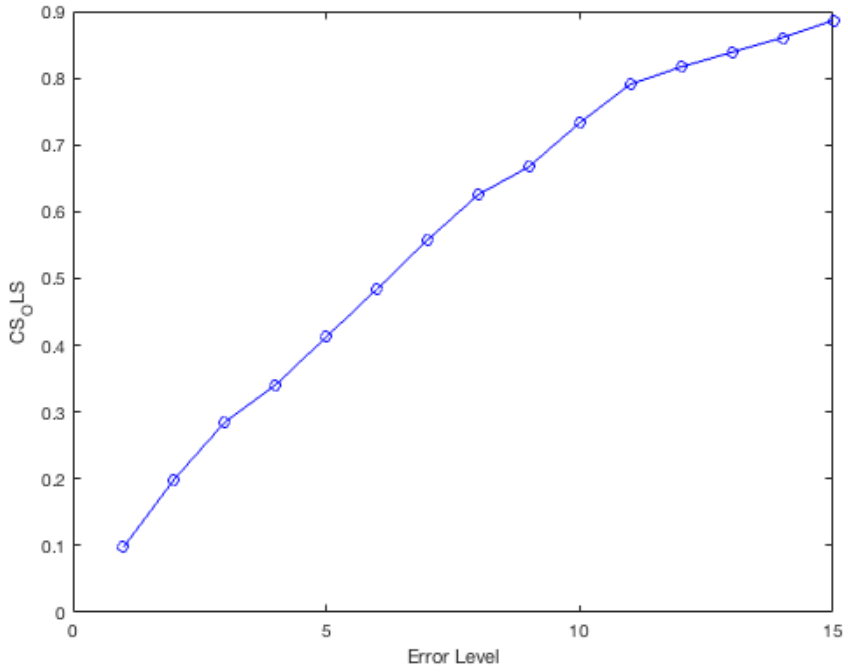
The results for Mean Absolute Error (MAE) and Cumulative Score (CS) for Linear Regression have been listed in the table below.

5. The cumulative error level has been set between 1 to 15 and a plot is generated of the CS value against the cumulative error level. The following code has been implemented.

```
%% generate a cumulative score (CS) vs. error level plot by varying the error level from 1 to 15.
for i = 1:15
    CS(i) = sum(error_lr < i == 1)/ size(ytest, 1);
end

plot (CS, 'b-o')

xlabel('Error Level')
ylabel('CS_OLS')
```



6. The MAE and CS (with cumulative error level 5) has been computed for both partial least square regression model and regression tree model using the Matlab built-in functions. The following code has been implemented:

```
%% Compute the MAE and CS value (with cumulative error level of 5) for both partial least square regression a

% Partial least square regression
[XL,yl,XS,YS,beta,PCTVAR,MSE,stats] = plsregress(xtrain, ytrain, 10);

yhat_test_plsr = [ones(size(xtest,1),1) xtest]*beta;

error_plsr = abs(yhat_test_plsr-ytest);
mae_plsr = sum(error_plsr)/size(ytest, 1);
i=5;
CS_plsr = sum(error_plsr < i == 1)/ size(ytest, 1);

fprintf('MAE(Partial Least square regression) = %f\n', mae_plsr);
fprintf('CS(5) = %f\n', CS_plsr);
```

```
%%

% Regression tree
w_rt = fitrtree(xtrain, ytrain);
yhat_test_rt = predict(w_rt, xtest);

error_rt = abs(yhat_test_rt-ytest);
mae_rt = sum(error_rt)/size(ytest, 1);
i=5;
CS_rt = sum(error_rt < i == 1)/ size(ytest, 1);

fprintf('MAE(regression tree) = %f\n', mae_rt);
fprintf('CS(5) = %f\n', CS_rt);
```

The results for Mean Absolute Error (MAE) and Cumulative Score (CS) for Partial Least Square Regression and Regression Tree have been listed in the table below.

7. The MAE and CS (with cumulative error level 5) has been computed for Support Vector Regression using the LIBSVM Toolbox. The following code has been implemented:

```
%% Compute the MAE and CS value (with cumulative error level of 5) for Support Vector Regression by using LIBSVM toolbox
addpath(genpath('/Users/tanmaiyiirao/Desktop/ecs797/lab1/ECS797Lab1/software'));
svr = fitrsvm(xtrain, ytrain);
yhat_test_svr = predict(svr, xtest);

error_svr = abs(yhat_test_svr-ytest);
mae_svr = sum(error_svr)/size(ytest, 1);
i=5;
CS_svr = sum(error_svr < i == 1)/ size(ytest, 1);

fprintf('MAE(SVR) = %f\n', mae_svr);
fprintf('CS(5) = %f\n', CS_svr);
```

The results for Mean Absolute Error (MAE) and Cumulative Score (CS) for Support Vector Regression have been listed in the table below.

8. Results for Error Level 5

Model	Mean Absolute Error (MAE)	Cumulative Score (CS)
Linear Regression (OLS)	7.704359	0.412351
Partial Least Square Regression (PSLR)	6.070294	0.525896
Regression Tree (RT)	8.176074	0.488048
Support Vector Regression (SVR)	5.731360	0.537849

Plot of Cumulative Scores (CS) for error level between 1 to 15 for the different models.