School of Electronic Engineering and Computer Science

ECS797 Machine Learning for Visual Data Analysis

Lab 3: Age Estimation by Regression

Name: Yogen Parekh

Student ID: 190306340

3. Complete the lab3.m file

1. data_age.mat file was loaded which contains the training and test data, and path was set for results using the given code. Training data has 500 instances with 201 features. Test data has 502 instances with 201 features.

2. This section of code calls built-in function regress() which takes numeric vectors xtrain and ytrain i.e. the training data and learns a Multiple Linear Regression model.

3. The testing section in lab3.m reads the testing data and applies the learned linear regression model to estimate the age for each input test data point.

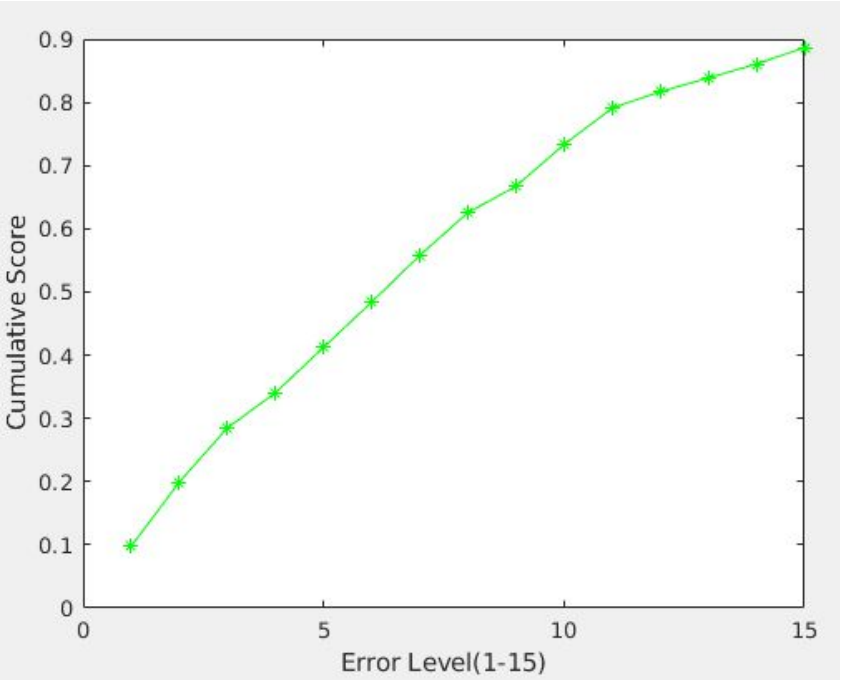
4. The Mean Absolute Error(MAE) and Cumulative Score(CS) with a cumulative error level of 5 has been computed in this code section by comparing the estimated ages with the ground truth ages. The corresponding snippet of the code and the resultant MAE and CS are attached here.

```
MAE Lin-Reg = 7.704359
CS Lin-Reg = 0.412351
```

```
all_errors = abs(yhat_test-ytest); % Computing all errors
mae_linreg = sum(all_errors)/size(ytest, 1);   % MAE
cs_linreg = sum(all_errors < err_level == 1)/ size(ytest, 1);  % CS

fprintf('MAE Lin-Reg = %f\n', mae_linreg);
fprintf('CS Lin-Reg = %f\n', cs_linreg);
```

5. A plot is generated here for the CS value against the corresponding Cumulative Error Level which varies from 1 to 15. The code used for this purpose and the required plot are attached below.

```
for i = 1:15
    % Computing CS for varying error level
    cs_linreg(i) = sum(all_errors < i == 1)/ size(ytest, 1);
end

plot (cs_linreg, 'g-*')

xlabel('Error Level(1-15)')
ylabel('Cumulative Score')
```



6. Using the following piece of code, I have computed the MAE and CS with the cumulative error level of 5 for both partial least square regression model and the regression tree model by using the Matlab built-in functions.

- **Partial Least Square Regression:**

```
MAE PLS-Reg = 6.070294
CS PLS-Reg = 0.525896
```

```
% Learning the PLS Regression model
[XL,YL,XS,YS,beta,PCTVAR,MSE,stats] = plsregress(xtrain, ytrain, 10);

 % Predicting the ages
yhat_test_p = [ones(size(xtest,1),1) xtest]*beta;

all_errors_p = abs(yhat_test_p-ytest); % Computing all errors
mae_p = sum(all_errors_p)/size(ytest, 1);   % MAE
cs_p = sum(all_errors_p < err_level == 1)/ size(ytest, 1);  % CS

fprintf('\n');
fprintf('MAE PLS-Reg = %f\n', mae_p);
fprintf('CS PLS-Reg = %f\n', cs_p);
```

- **Regression Tree:**

```
MAE Reg-Tree = 8.235005
CS Reg-Tree = 0.496016
```

```
% Learning the Regression tree model
w_rt = fitrtree(xtrain, ytrain);
yhat_test_rt = predict(w_rt, xtest); % Predicting the ages

all_errors_rt = abs(yhat_test_rt-ytest); % Computing all errors
mae_rt = sum(all_errors_rt)/size(ytest, 1);   % MAE
cs_rt = sum(all_errors_rt < err_level == 1)/ size(ytest, 1);  % CS

fprintf('\n');
fprintf('MAE Reg-Tree = %f\n', mae_rt);
fprintf('CS Reg-Tree = %f\n', cs_rt);
```

7. This section computes the MAE and CS values with cumulative error level of 5 for Support VectorRegression(SVR) using the LIBSVM toolbox provided at the specified link. The code is attached below.

```
MAE SV-Reg = 5.623867
CS SV-Reg = 0.565737
```

```
addpath(genpath('./libsvm'));

% Fiting model with command line parameters
svr = svmtrain(ytrain, xtrain, '-s 3 -t 0');
yhat_test_svr = svmpredict(ytest, xtest, svr); % Predicting the ages

all_errors_svr = abs(yhat_test_svr-ytest); % Computing all errors
mae_svr = sum(all_errors_svr)/size(ytest, 1);   % MAE
cs_svr = sum(all_errors_svr < err_level == 1)/ size(ytest, 1);  % CS

fprintf('\n');
fprintf('MAE SV-Reg = %f\n', mae_svr);
fprintf('CS SV-Reg = %f\n', cs_svr);
```

- Comparing the Mean Absolute Errors(MAE) and Cumulative Score(CS) at error level 5 from all the four models:

Model	Mean Absolute Error (MAE)	Cumulative Score (CS)
Linear Regression	7.704359	0.412351
Partial Least Square Regression	6.070294	0.525896
Regression Tree	8.235005	0.496016
Support Vector Regression	5.623867	0.565737