

CS 210 Project Blog

MENU

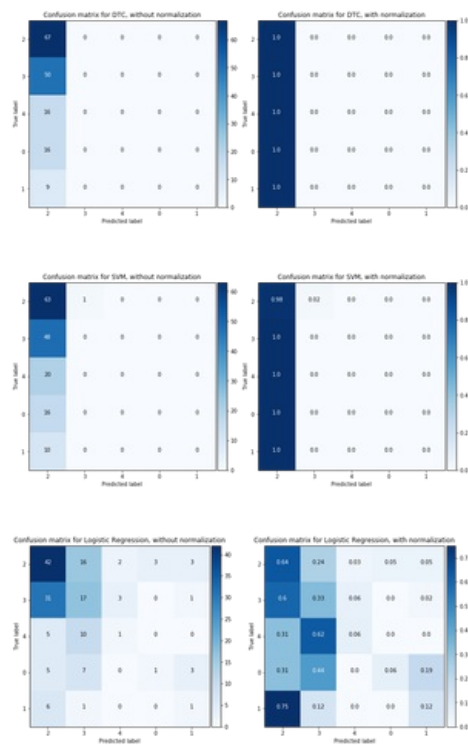
Confusion Matrices

There are the confusion matrices for all cases.

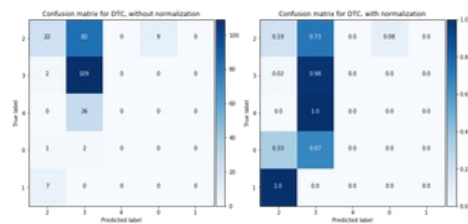
(Note that the numbers may be differ from cross-validated numbers.)

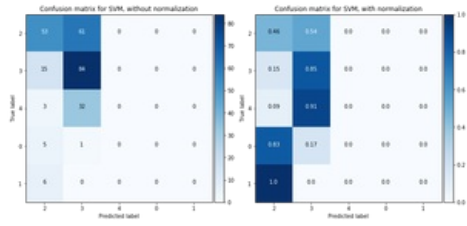
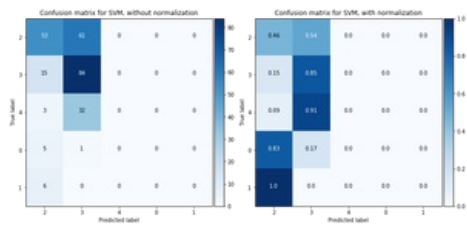
For no test scores given:

for Mat:



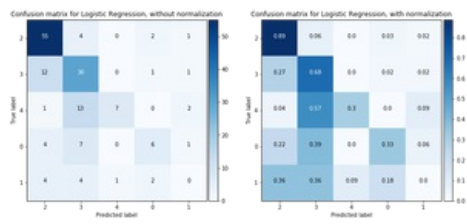
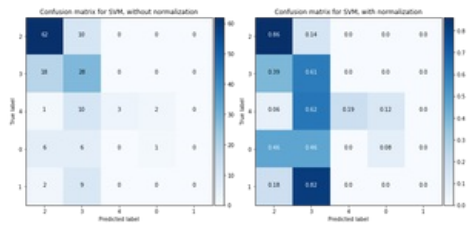
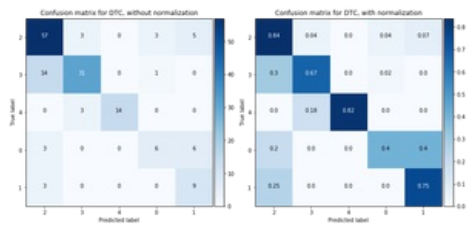
for Por:



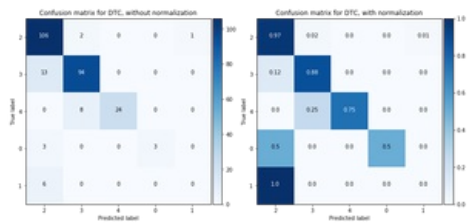


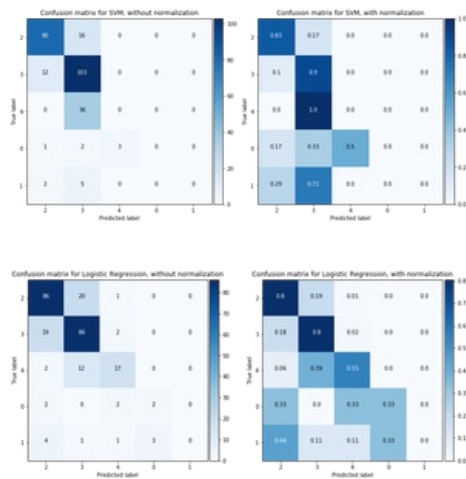
For both test scores given:

for Mat:



for Por:





May 3, 2017 / Leave a comment

Comparison of Different Models

DBelow, you can find the comparison of OneR, Decision Tree, SVM, Logistic Regression and Linear Regression for my data:

None Given	One R	Decision Tree	SVM	Logistic Regression		Linear Regression
Mat	0.41	0.41	0.42	0.39		—
Por	0.49	0.52	0.53	0.55		—

Both Given	One R	Decision Tree	SVM	Logistic Regression		Linear Regression
Mat	0.63	0.78	0.60	0.62		0.82
Por	0.75	0.85	0.77	0.69		0.85

April 29, 2017 / Leave a comment

Logistic Regressions

I've classified the data using Logistic Regression.

Below are the accuracies:

for Mat:

No test scores: 0.39

Both test scores: 0.62

for Por:

No test scores: 0.55

Both test scores: 0.69

You can find the confusion matrices in the code.

Codes: [MatLogisticRegression.pdf](#) [PorLogisticRegression.pdf](#)

April 29, 2017 / [Leave a comment](#)

SVMs

I've classified the data using SVM.

Below are the accuracies:

for Mat:

No test scores: 0.42

Both test scores: 0.60

for Por:

No test scores: 0.53

Both test scores: 0.77

You can find the confusion matrices in the code.

Codes: [MatSVM.pdf](#) [PorSVM.pdf](#)

April 29, 2017 / [Leave a comment](#)

Decision Trees

I've constructed decision trees for both no test scores given and both test scores are given.

The results are as follows:

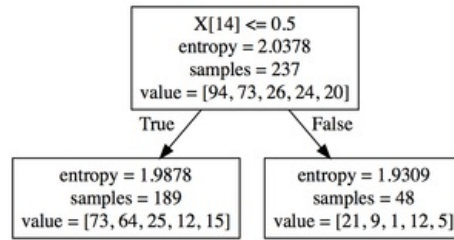
No test scores are given:

for Mat:

Accuracy: 0.41

Best: MaxDepth = 1 (means one R)

Decision Tree:

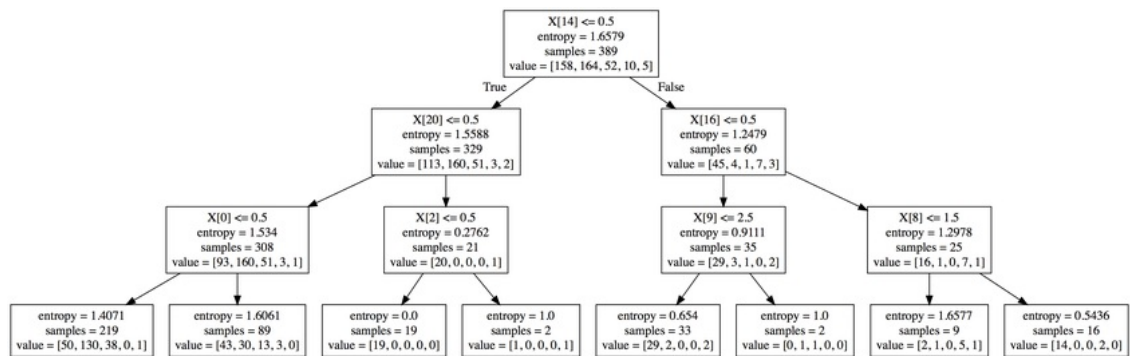


for Por:

Accuracy: 0.52

Best: MaxDepth = 3

Decision Tree:



Both test scores are given:

for Mat:

Accuracy: 0.78

Best: MaxDepth = 4

Decision Tree:

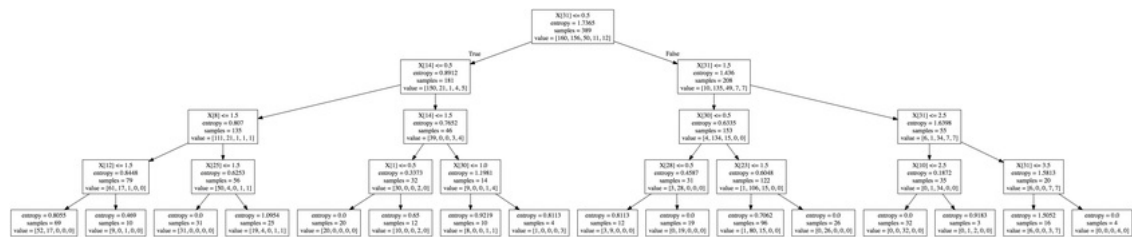


for Por:

Accuracy: 0.85

Best: MaxDepth = 4

Decision Tree:



You can find confusion matrices in the [code.pdf](#).

Code: [MatDecisionTree.pdf](#) [PorDecisionTree.pdf](#)

April 29, 2017 / Leave a comment

Linear Regressions

I've performed linear regression on (numerical) data.

For Mat, R-squared: 0.822

For Por, R-squared: 0.848

They are very good results and mean that we can very accurately estimate G3 given G1 and G2.

Code (for Mat): [MatLinearRegression.pdf](#)

Code (for Por): [PorLinearRegression.pdf](#)

April 29, 2017 / Leave a comment

Getting and Processing Data

I've downloaded the data and convert some numeric data to categorical through binning.

I've binned the numerical data as follows:

- age:
 - age <= 18
 - age > 18
- absences
 - absences <= 20
 - 20 < absences <= 40

- $40 < \text{absences} \leq 60$
- $60 < \text{absences} \leq 80$
- $80 < \text{absences}$
- G1, G2, G3
 - $0 \leq \text{score} < 4$
 - $4 \leq \text{score} < 8$
 - $8 \leq \text{score} < 12$
 - $12 \leq \text{score} < 16$
 - $16 \leq \text{score} \leq 20$

Code: [CategorizeData.pdf](#)

April 29, 2017 / [Leave a comment](#)



RECENT POSTS

- [Reports](#)
May 19, 2017
- [ZeroR Update](#)
May 19, 2017
- [Equal Weight DTCs](#)
May 14, 2017
- [Plan for last phase](#)
May 12, 2017
- [Logistic Regressions for Percentages](#)
May 7, 2017
- [Decision Trees for Percentages](#)
May 7, 2017
- [Processing Data for Percentages](#)
May 7, 2017
- [Confusion Matrices](#)
May 3, 2017
- [Comparison of Different Models](#)
April 29, 2017
- [Logistic Regressions](#)
April 29, 2017
- [SVMs](#)
April 29, 2017
- [Decision Trees](#)
April 29, 2017
- [Linear Regressions](#)
April 29, 2017

- [Getting and Processing Data](#)
April 29, 2017
- [GitHub Integration](#)
April 26, 2017
- [Part 1 Overview](#)
April 24, 2017
- [Planning for the rest of the project](#)
April 24, 2017
- [Regression \(New\)](#)
April 14, 2017
- [Cluster](#)
April 13, 2017
- [Regressions and Some Plots](#)
April 9, 2017
- [Minimized Data](#)
April 8, 2017
- [Correlation Matrix & Scatter Plot Matrix](#)
March 24, 2017
- [Filtered Missing Data](#)
March 24, 2017
- [Merged Data](#)
March 23, 2017
- [Code for Histograms](#)
March 18, 2017
- [Histograms](#)
March 18, 2017
- [Data Downloaded](#)
March 18, 2017
- [Project Proposal \(v3\)](#)
March 18, 2017
- [Project Proposal \(v2\)](#)
March 2, 2017
- [Project Proposal](#)
March 1, 2017