Juan Pedro López Cabrera, 12-Nov-2015

RoboticaUVic-PR-week3

UVic Robotics Master. Pattern Recognition Homework 1

Solving Q2:
theta = [22.53280632]
MSE = 84.4195561562

Solving Q3:
CRIM
MSE training set: 65.7733954
MSE testing set: 873.51687362
R2 of testing set against theta from training set: -8.3637876

ZN
MSE training set: 62.18486413
MSE testing set: 91.93379466
R2 of testing set against theta from training set: 0.01450269

INDUS
MSE training set: 60.2029463
MSE testing set: 74.74611197
R2 of testing set against theta from training set: 0.19874849

CHAS
MSE training set: 69.24037547
MSE testing set: 105.20352304
R2 of testing set against theta from training set: -0.12774404

NOX
MSE training set: 61.54527906
MSE testing set: 80.16012623
R2 of testing set against theta from training set: 0.1407122

RM
MSE training set: 15.93514792
MSE testing set: 78.05563443
R2 of testing set against theta from training set: 0.16327159

AGE
MSE training set: 61.42749592
MSE testing set: 88.95536072

R2 of testing set against theta from training set: 0.04643044

DIS
MSE training set: 68.00062308
MSE testing set: 97.46324219
R2 of testing set against theta from training set: -0.044771

RAD
MSE training set: 69.14935938
MSE testing set: 144.90523085
R2 of testing set against theta from training set: -0.5533321

TAX
MSE training set: 65.4358816
MSE testing set: 67.51748098
R2 of testing set against theta from training set: 0.27623682

PTRATIO
MSE training set: 59.75097258
MSE testing set: 72.50716333
R2 of testing set against theta from training set: 0.22274921

B
MSE training set: 66.04075277
MSE testing set: 86.44950014
R2 of testing set against theta from training set: 0.07329236

LSTAT
MSE training set: 34.38075578
MSE testing set: 43.3437236
R2 of testing set against theta from training set: 0.53537083

Most informative: LSTAT, because it produces the best MSE
Best generalization: TAX, because MSE on both testing and training are the most similar
Worst generalization: CRIM, becase it produces the worst MSE

Solving Q4:
With CRIM
MSE training set: 9.98751733
MSE testing set: 303.43686293
R2 of testing set against theta from training set: -2.25273434

Without CRIM
MSE training set: 10.21628339
MSE testing set: 50.49037689
R2 of testing set against theta from training set: 0.45876127

After removing CRIM, which looked the worst-performing variable, the MSE is reduced drastically and R2 takes a better looking value in the range 0,1

Solving Q5:

Adding more capacity to this model seems to improve some of the variables while producing worst results for others.

CRIM
MSE training set: 62.29064434
MSE testing set: 5.00715676e+13
R2 of testing set against theta from training set: -5.36749247e+11

ZN
MSE training set: 59.94340249
MSE testing set: 103.84142845
R2 of testing set against theta from training set: -0.11314287

INDUS
MSE training set: 49.82500558
MSE testing set: 87.51560776
R2 of testing set against theta from training set: 0.06186407

CHAS
MSE training set: 69.24037992
MSE testing set: 105.21129762
R2 of testing set against theta from training set: -0.12782738

NOX
MSE training set: 61.14210648
MSE testing set: 102.94522166
R2 of testing set against theta from training set: -0.10353585

RM
MSE training set: 13.32255635
MSE testing set: 80.08419049
R2 of testing set against theta from training set: 0.1415262

AGE
MSE training set: 59.25425217
MSE testing set: 87.72147221
R2 of testing set against theta from training set: 0.05965728

DIS
MSE training set: 63.57334658
MSE testing set: 116.31840985
R2 of testing set against theta from training set: -0.24689164

RAD
MSE training set: 62.7858572

MSE testing set: 720635.34520195
R2 of testing set against theta from training set: -7723.95244339

TAX
MSE training set: 63.3427603
MSE testing set: 228316.94988679
R2 of testing set against theta from training set: -2446.4758165

PTRATIO
MSE training set: 55.60500804
MSE testing set: 68.09664278
R2 of testing set against theta from training set: 0.27002841

B
MSE training set: 65.35835443
MSE testing set: 126.01732834
R2 of testing set against theta from training set: -0.35086056

LSTAT
MSE training set: 19.78887402
MSE testing set: 40.99735048
R2 of testing set against theta from training set: 0.56052311