06 MAYIS 2021

CSE-315

MIDTHERM REPORT

FERHAT KAMALI
171805056

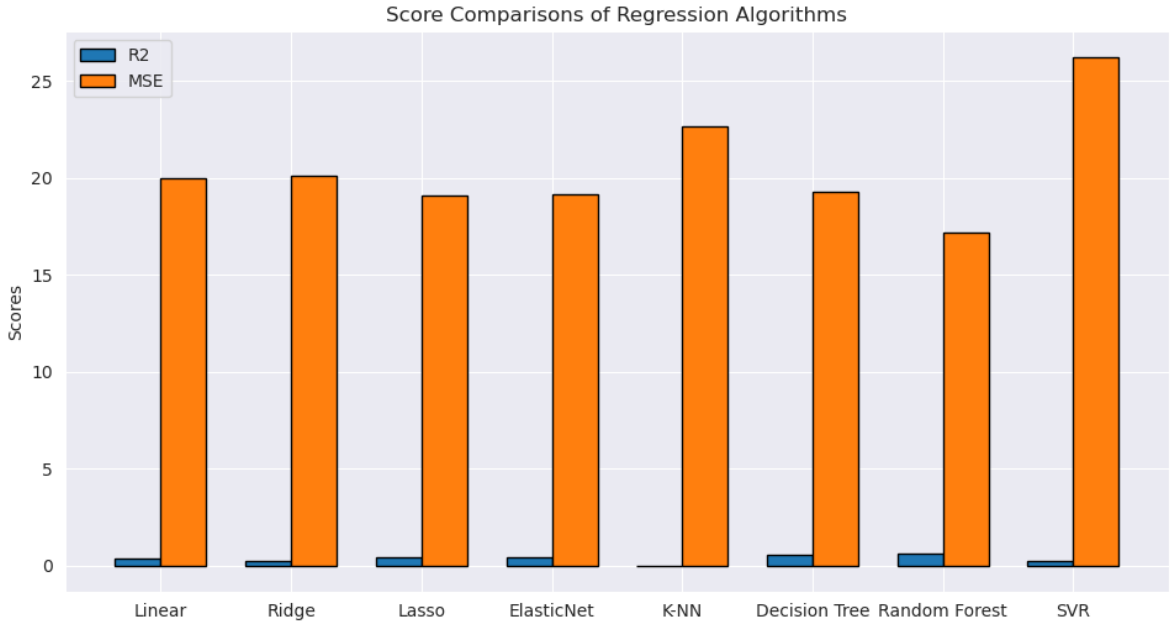
ANIL DURMAZ
181805080

MUHAMMED FURKAN YILMAZ
181805069

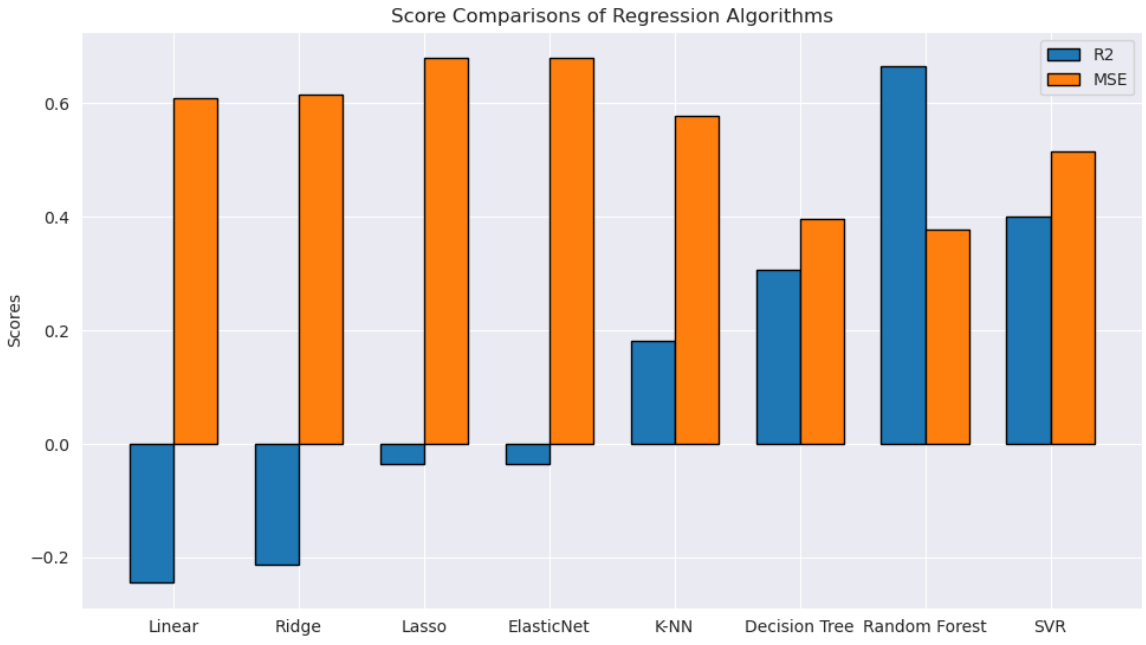
1. How using Preprocessed data instead of the raw data affected the results?

Preprocessed data has much better results than raw data in case of R2 values.

No-preprocessed (raw) data:



Preprocessed data:



2. What kind of learning performance you gained by using preprocessed data?

R2 results with Raw data:

Linear 0.391103

Ridge 0.357409

Lasso 0.344928

ElasticNet 0.333247

K-NN 0.124199

Decision Tree 0.745353

Random Forest 0.628524

SVR 0.205959

R2 results with prerpocessed data:

Linear -0.060427

Ridge -0.215622

Lasso -0.035490

ElasticNet -0.035490

K-NN 0.128493

Decision Tree 0.636484

Random Forest 0.592557

SVR 0.397328

According to these results we can see that preprocessed data highly increases the predictions' truth possibility in case of R2.

3. After all the processes, which algorithm has the better learning?

Linear

R2 = 0.18232580849001068

MSE = 0.46682120308563096

Ridge

R2 = 0.18192235736642393

MSE = 0.4670515387631594

Lasso

R2 = -0.002427503486010396

MSE = 0.5722993559565563

ElasticNet

R2 = -0.002427503486010396

MSE = 0.5722993559565563

KKN

R2 = 0.2421052631578947

MSE = 0.4326923076923077

DecisionTree

R2 = 0.848421052631579

MSE = 0.08653846153846154

RandomForest

R2 = 0.7978947368421052

MSE = 0.11538461538461539

SVR

R2 = 0.2774800291621552

MSE = 0.41249637758170227

Linear 0.391103

Ridge 0.357409

Lasso 0.344928

ElasticNet 0.333247

K-NN 0.124199

Decision Tree 0.745353

Random Forest 0.628524

SVR 0.205959

prepros

Linear 0.060427

Ridge 0.215622

Lasso 0.035490

ElasticNet 0.035490

K-NN 0.128493

Decision Tree 0.636484

Random Forest 0.592557

SVR 0.397328

After all the processes, for R2 best algorithm is Decision Tree, but for MSE best algorithms are ElasticNet and lasso.