ML Project Document

Numercal :

a_ info on dataset :

name of the dataset is Used-Cars-Catalog

number of classes is regression task (predicting the used car price)

total number of samples in dataset is 37000

number of samples used in training is 30000& testing is 7000

b_implementation details :

There were x features extracted and they are transmission,odometer_value,engine_fuel,engine_capacity,drivetrain,price_usd,num_of_photos,up_counter,duration_listed,age

yes cross-validation is used in linear regression and the number of fold is 5 and the ratio of training/validation is

c_Results details :

Accuracy :

```
Linear model accurcy score : 0.6612307585923843
Random Forest Regression accurcy score : 0.8675919770045515
```

Image :

a_ info on dataset :

name of the dataset is age estimation

number of classes is unsupervised (clustring task)

total number of samples in dataset is 3252

number of samples used in training is 2601 & testing is 650

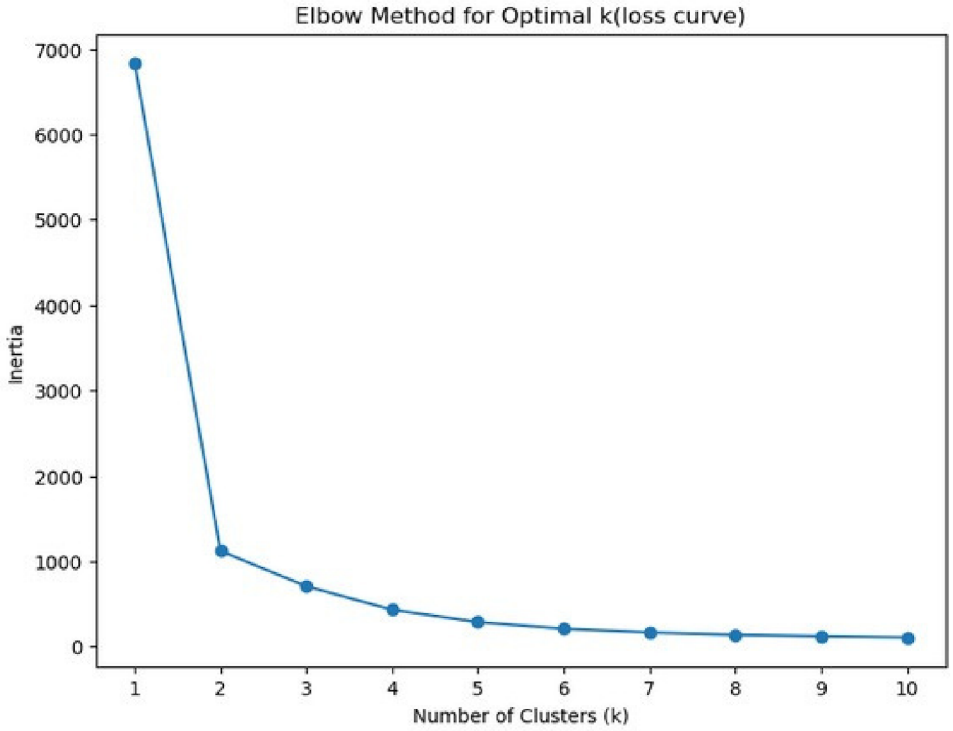
b_implementation details :

There were 3 features extracted and they are age and race and date&time

No cross-validation is not used

c_Results details :

loss accuracy :

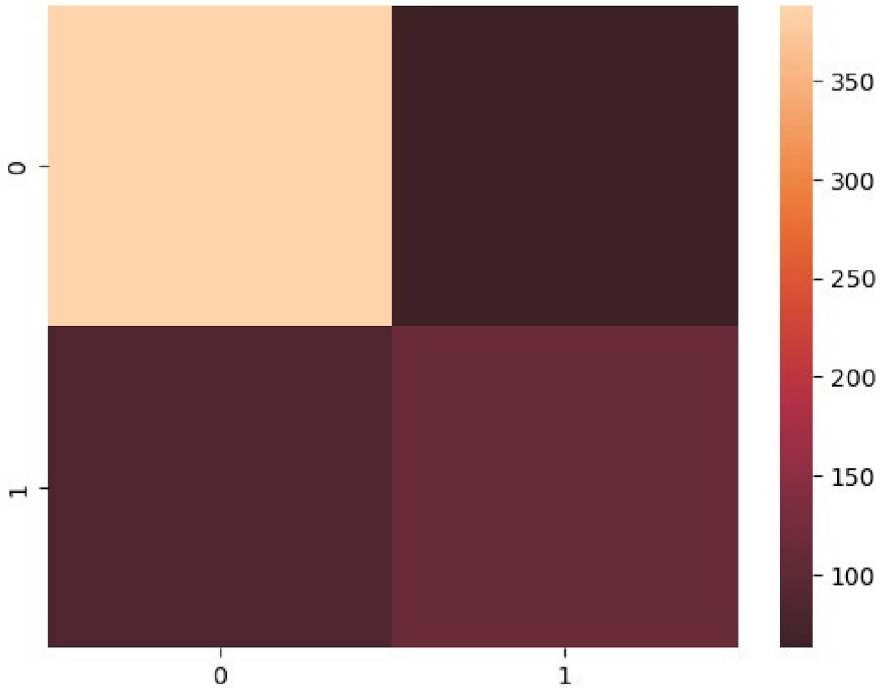


accuracy : 0.7649

confusion metrics :

338 63

90 110



ROC Curve :