

MVP

# Mohammed A. Saeedi

---

MVP

[msaeedi@mrda.gov.sa](mailto:msaeedi@mrda.gov.sa)

11/11/2021

## My data-set details:

Shape is (581012, 55)

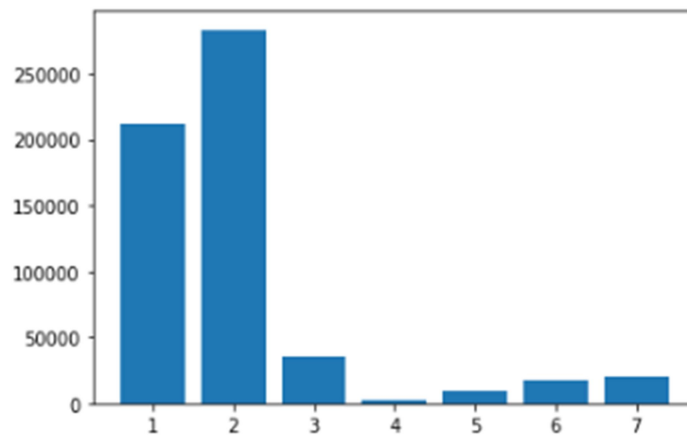
Number of instances = 581012

Number of attributes = 55

## Class information:

Class	Number of rows
2	283301
1	211840
3	35754
7	20510
6	17367
5	9493
4	2747

Each class number mean type of cover type.



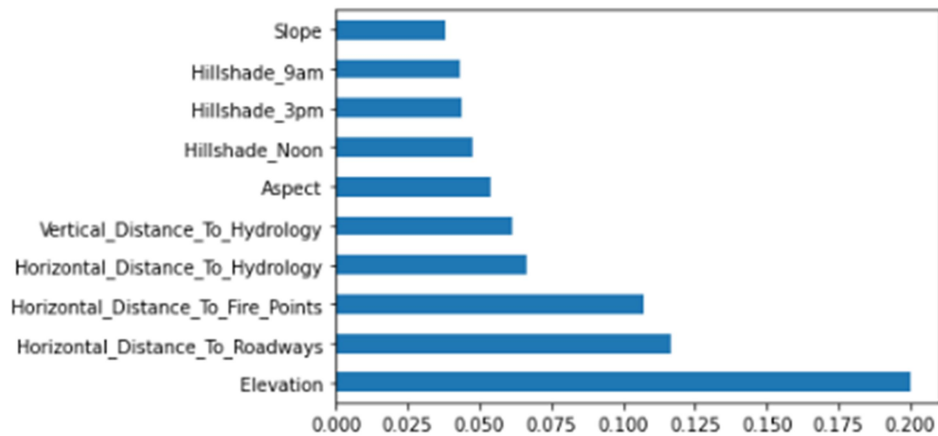
Number of missing value = 0

Number of duplicate rows = 0

I converted from multi class to binary class so my data-set after converted

class	Number of rows
0	297711
1	283301

My features selection technique I went with ExtraTreesClassifier i took the 10 important features, I plot graph of feature importance for better visualization.



I spilt the data set into train – validation – test

Number of instances in the training dataset = 325366

Number of instances in the validation dataset = 139443

Number of instances in the test dataset = 116203

My model will predict if the input data class 1 or not.

I will use three model:

- Decision Tree Classifier
- Naïve bayes
- Random Forest Classifier

I will show accuracy and confusion matrix for each classifier.

Then I will apply 10-Fold Cross Validation for all classifier.

Finally I will compare the results of each classifier.