Univariate Selection

```
import pandas as pd
import numpy as np
from sklearn.feature selection import SelectKBest
from sklearn.feature selection import chi2
data = pd.read csv("/content/drive/MyDrive/mobile train.csv")
X = data.iloc[:,0:20] #independent columns
y = data.iloc[:,-1] #target column i.e price range
#apply SelectKBest class to extract top 10 best features
bestfeatures = SelectKBest(score func=chi2, k=10)
fit = bestfeatures.fit(X,y)
dfscores = pd.DataFrame(fit.scores )
dfcolumns = pd.DataFrame(X.columns)
#concat two dataframes for better visualization
featureScores = pd.concat([dfcolumns,dfscores],axis=1)
featureScores.columns = ['Specs','Score'] #naming the dataframe columns
```

featureScores

	Specs	Score	7 -
0	battery_power	14129.866576	
1	blue	0.723232	
2	clock_speed	0.648366	
3	dual_sim	0.631011	
4	fc	10.135166	
5	four_g	1.521572	
6	int_memory	89.839124	
7	m_dep	0.745820	

print(featureScores.nlargest(10,'Score')) #print 10 best features

	Specs	Score
13	ram	931267.519053
11	px_height	17363.569536
0	battery_power	14129.866576
12	px_width	9810.586750
8	mobile_wt	95.972863
6	int_memory	89.839124

```
15 sc_w 16.480319
16 talk_time 13.236400
4 fc 10.135166
14 sc_h 9.614878
```

Feature Importance

_

```
from sklearn.ensemble import ExtraTreesClassifier
import matplotlib.pyplot as plt
model = ExtraTreesClassifier()
model.fit(X,y)
```

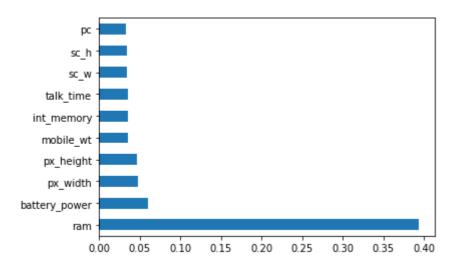
```
ExtraTreesClassifier()
```

print(model.feature_importances_) #use inbuilt class feature_importances of tree based cla

```
[0.0601292 0.02121841 0.03303608 0.01968605 0.03326206 0.01766871 0.03546954 0.03270203 0.03576236 0.03246003 0.03347872 0.04659445 0.04801884 0.39339916 0.03406896 0.03435857 0.03542379 0.0144744 0.0184762 0.02031244]
```

#plot graph of feature importances for better visualization

```
feat_importances = pd.Series(model.feature_importances_, index=X.columns)
feat_importances.nlargest(10).plot(kind='barh')
plt.show()
```



Correlation Matrix with Heatmap

import seaborn as sns
#get correlations of each features in dataset
corrmat = data.corr()

```
top_corr_features = corrmat.index
plt.figure(figsize=(20,20))
#plot heat map
g=sns.heatmap(data[top_corr_features].corr(),annot=True,cmap="RdYlGn")
```

battery_power -	1	0.011	0.011	-0.042		0.016	-0.004	0.034	0.0018	-0.03		0.015	-0.0084	-0.00065	-0.03	-0.021		0.012	-0.011	-0.0083	0.2
blue -	0.011				0.0036	0.013	0.041	0.004	-0.0086		-0.01	-0.0069	-0.042		-0.003	0.00061	0.014	-0.03	0.01	-0.022	0.021
clock_speed -	0.011			-0.0013	-0.00043	-0.043	0.0065	-0.014	0.012	-0.0057	-0.0052	-0.015	-0.0095	0.0034	-0.029	-0.0074	-0.011	-0.046		-0.024	-0.0066
dual_sim -	-0.042		-0.0013		-0.029	0.0032	-0.016	-0.022	-0.009	-0.025	-0.017	-0.021	0.014	0.041	-0.012	-0.017	-0.039	-0.014	-0.017	0.023	0.017
fc -	0.033	0.0036	-0.00043	-0.029		-0.017	-0.029	-0.0018	0.024	-0.013	0.64	-0.01	-0.0052	0.015	-0.011	-0.012	-0.0068	0.0018	-0.015		0.022
four_g -	0.016	0.013	-0.043	0.0032	-0.017		0.0087	-0.0018	-0.017	-0.03	-0.0056	-0.019	0.0074	0.0073			-0.047	0.58	0.017	-0.018	0.015
int_memory -	-0.004	0.041	0.0065	-0.016	-0.029	0.0087		0.0069	-0.034	-0.028	-0.033	0.01	-0.0083			0.012	-0.0028	-0.0094	-0.027	0.007	0.044
m_dep -	0.034	0.004	-0.014	-0.022	-0.0018	-0.0018	0.0069			-0.0035		0.025	0.024	-0.0094	-0.025	-0.018	0.017	-0.012	-0.0026	-0.028	0.00085
mobile_wt -	0.0018	-0.0086	0.012	-0.009	0.024	-0.017	-0.034			-0.019		0.00094	9e-05	-0.0026	-0.034	-0.021	0.0062	0.0016	-0.014	-0.00041	-0.03
n_cores -	-0.03		-0.0057	-0.025	-0.013	-0.03	-0.028	-0.0035	-0.019		-0.0012	-0.0069	0.024	0.0049	-0.00031	0.026	0.013	-0.015	0.024	-0.01	0.0044
pc -	0.031	-0.01	-0.0052	-0.017	0.64	-0.0056	-0.033	0.026		-0.0012		-0.018	0.0042		0.0049	-0.024	0.015	-0.0013	-0.0087	0.0054	0.034
px_height -	0.015	-0.0069	-0.015	-0.021	-0.01	-0.019	0.01	0.025	0.00094	-0.0069	-0.018		0.51	-0.02		0.043	-0.011	-0.031	0.022		0.15
px_width -	-0.0084	-0.042	-0.0095	0.014	-0.0052	0.0074	-0.0083	0.024	9e-05	0.024	0.0042	0.51	1	0.0041	0.022	0.035	0.0067	0.00035	-0.0016		0.17
ram -	-0.00065	0.026	0.0034	0.041	0.015	0.0073		-0.0094	-0.0026	0.0049	0.029	-0.02	0.0041		0.016		0.011		-0.03		0.92
sc_h -	-0.03	-0.003	-0.029	-0.012	-0.011			-0.025	-0.034	-0.00031	0.0049		0.022			0.51	-0.017	0.012	-0.02		0.023
sc_w -	-0.021	0.00061	-0.0074	-0.017	-0.012		0.012	-0.018	-0.021	0.026	-0.024	0.043			0.51	1	-0.023	0.031	0.013		0.039

- 0.8

- 0.6

- 0.4

0

talk_time -	0.053	0.014	-0.011	-0.039	-0.0068	-0.047	-0.0028	0.017	0.0062	0.013	0.015	-0.011	0.0067	0.011	-0.017	-0.023	1	-0.043	0.017	-0.03	0.022
three_g -	0.012	-0.03	-0.046	-0.014	0.0018	0.58	-0.0094	-0.012	0.0016	-0.015	-0.0013	-0.031	0.00035	0.016	0.012		-0.043		0.014	0.0043	0.024
touch_screen -	-0.011	0.01		-0.017	-0.015	0.017	-0.027	-0.0026	-0.014	0.024	-0.0087	0.022	-0.0016	-0.03	-0.02	0.013	0.017	0.014		0.012	-0.03
wifi -	-0.0083	-0.022	-0.024	0.023		-0.018	0.007	-0.028	-0.00041	-0.01	0.0054			0.023	0.026		-0.03	0.0043	0.012		0.019
price_range -		0.021	-0.0066	0.017	0.022		0.044	0.00085	-0.03	0.0044	0.034							0.024	-0.03		1
	battery_power -	plue -	dock_speed -	dual_sim -	ft -	four_g -	int_memory -	m_dep_	mobile_wt -	n_cores -	Ж.	px_height -	px_width -	ram -	- u¯x	- W_X	talk_time -	three_g -	touch_screen -	wifi -	price_range -

- 0.0