CAPSTONE PROJECT (AD CLICK PREDICTION DATASET)

The objective of this dataset is to predict whether a user will click on an online ad based on their demographics, browsing behaviour, the context of the ad's display, and the time of day.

MODEL TESTING:

We will need to clean the data and then apply machine learning models to predict and evaluate data. It is a really challenging request for this kind of data. Select K Best classification we using the K best Value

K best =5

	Logistic	SVMl	SVMnl	KNN	Navie	Decision	Random
ChiSquare	0.61039	0.61039	0.614209	0.754775	0.613445	0.912911	0.896104

K best =4

	Logistic	SVMl	SVMnl	KNN	Navie	Decision	Random
ChiSquare	0.61039	0.61039	0.611917	0.781513	0.613445	0.921314	0.919786

K best =3

	Logistic	SVMl	SVMnl	KNN	Navie	Decision	Random
ChiSquare	0.61039	0.61039	0.611154	0.808251	0.611917	0.927426	0.92437

Then we calculate to prediction using many algorithms to feature selection method for ad click prediction. Finally, we find out the good algorithm for ad click prediction is Random Forest classification.

Accuracy score- It turns out that the accuracy score in this Random Forest algorithm is very good, which is equal to 0.81 comparatively with others algorithms. Overall performance of model.

Recall- value of 0.93 only correctly classified values get.

Precision- correct and wrong classified of class value is 0.80

We find out Recall and Precision always high is a good model, so we check f1 measure.

F1 score-0.86 model validation depends on the recall and precision value.

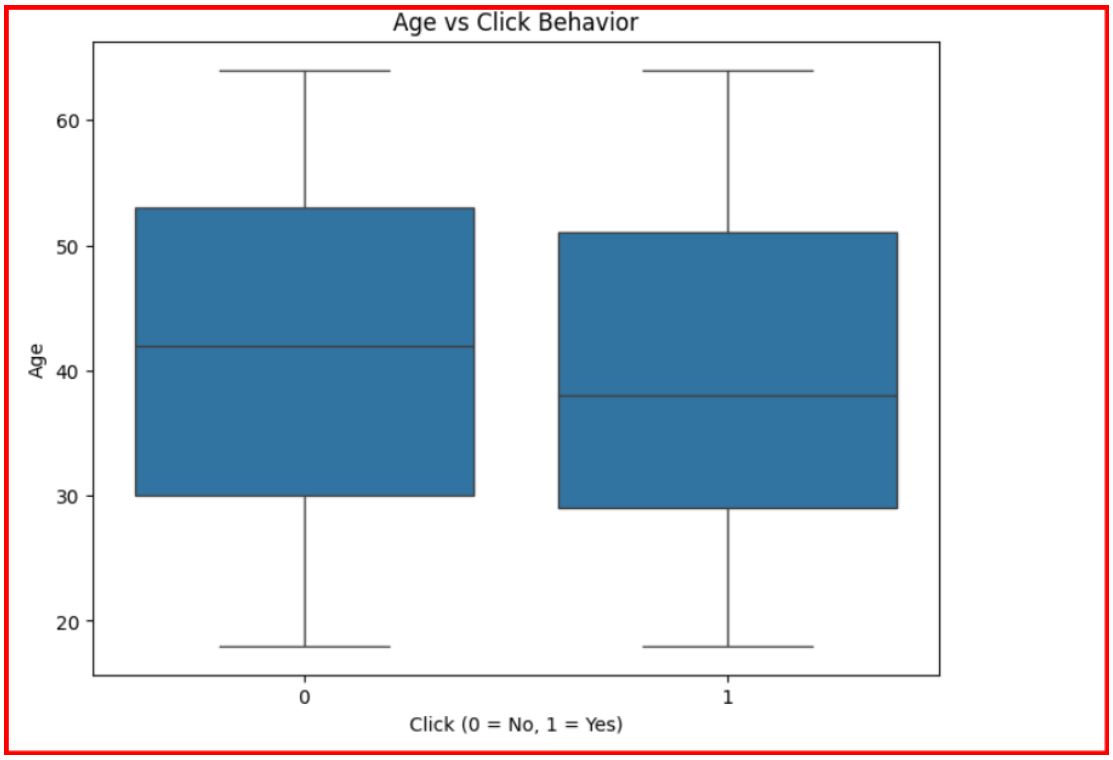
Receiver operating characteristic- area under curved (roc-auc)- we predict 0.93 almost 1 is a good model performance. We find grid prediction also in (gini, log2, max features) criterion.

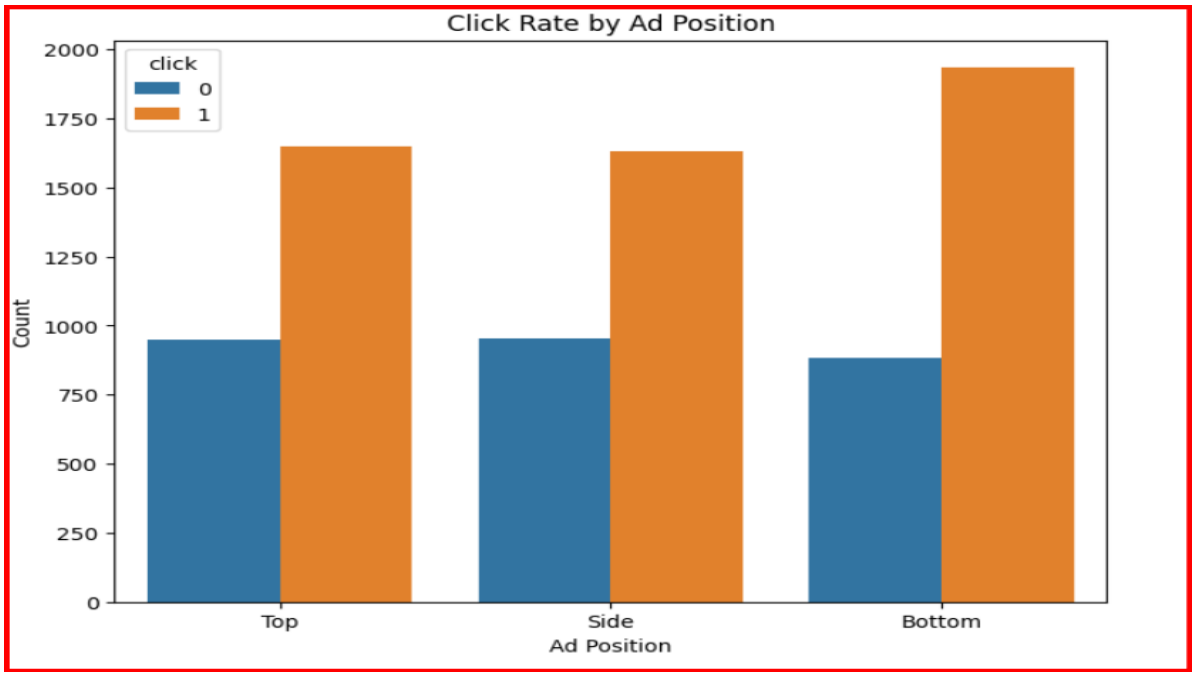
Finally, we come to the model testing stage by making predictions based on existing features.

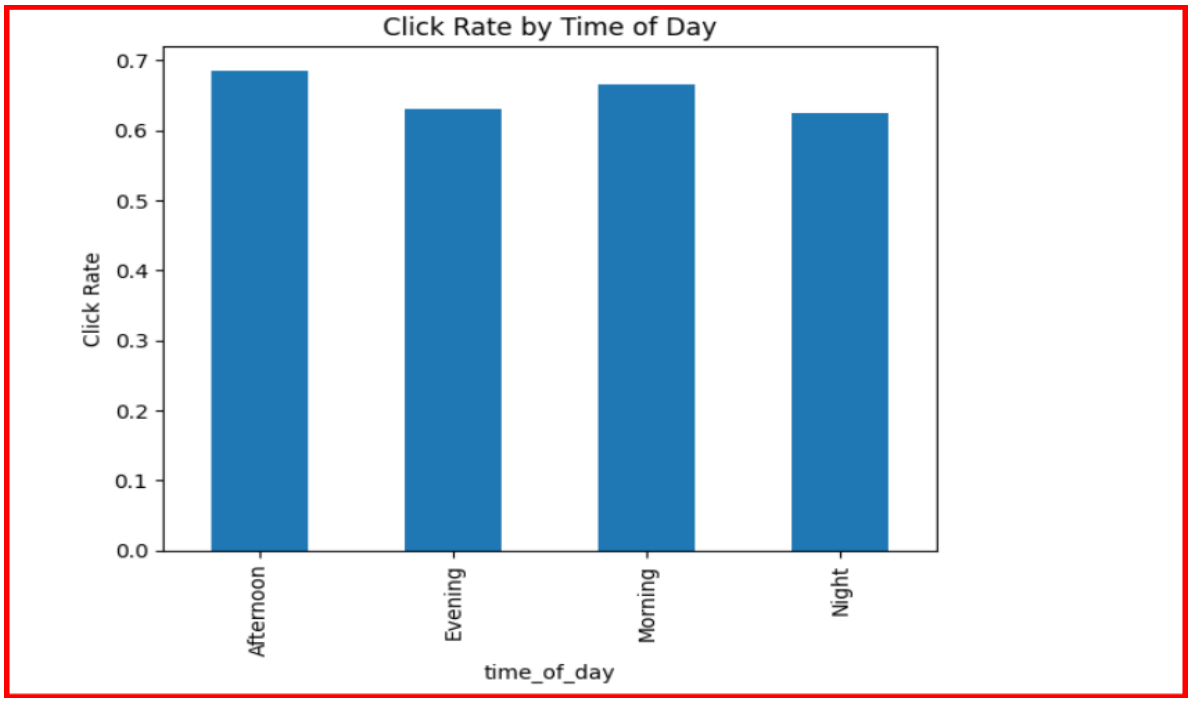
CONCLUSION:

We analyzing the dataset is Ad click prediction , through the daily rate based on time spent on the website like desktop device comparatively mobile & tablet seeing the ad and click . Next the position of Ad seeing people like bottom of the ad .Dataset of male & female equally seeing the ad and afternoon time more over ad click we analyzed the dataset.

This data can be used to improve ad targeting strategies, optimize ad placement, and better understand user interaction with online advertisements.




Click Rate by Ad Position
click
0
1
2000
1750
1500
1250
1000
750
500
250
0
Count
Top
Side
Bottom
Ad Position


Click Rate by Time of Day
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Click Rate
Afternoon
Evening
Morning
Night
time_of_day