

PROJECT REPORT

OVERVIEW:

Topic of the project is “**discount coupon application**” . In this project we check whether the given person details are applicable for a discount or not .It is a binary classification model .

Data was given in different files so first they were merged on the basis of the common attributes . Data Wrangling was performed on the features which were not contributing much to the result as well as on removing the string type features to boolean features.

In this project we used the `logisticRegression()` classifier with customizing its attributes to overcome the **problem of underfitting of data**. Accuracy was found to be 69% while the `roc_auc_score` was 75%.

```
In [30]: ▶ pre=clf.predict(x_test)
          acc=accuracy_score(pre,y_test)
          print(acc)
```

```
0.6886626124284546
```

```
In [31]: ▶ confusion_matrix(y_test,pre)
```

```
Out[31]: array([[95290, 46280],
               [ 2458, 12516]], dtype=int64)
```

```
In [32]: ▶ roc_auc_score(y_test,pre)
```

```
Out[32]: 0.7544716933900609
```
