


LOGISTIC REGRESSION REPORT

1. ASSUMPTIONS VERIFICATION

1.1. Assumption of Multicollinearity

These are the VIF values obtained for each and every columns, from these set of values it is clear that other than the age column rest of the columns are having VIF values between the range 1-10, VIF of age is 13.256 which is not a bad value, this value is almost close to 10.

The assumption of Multicollinearity is thus satisfied. In this step other than the "ID" feature I am not removing any other feature.

```
 vif=[]  
for i in range(X.shape[1]):  
    vif.append(variance_inflation_factor(X.values,i))  
  
pd.DataFrame({"Feature":X.columns,"VIF":vif}).sort_values(by="VIF",ascending=False)
```

Code Snippet 1.1 . VIF calculation python code

	Feature	VIF
14	age	13.256224
1	AMT_INCOME_TOTAL	6.170903
8	JOB_CATEGORY	4.063915
10	STATUS_OF_LOAN	3.775648
13	OWNS_PROPERTY	3.195439
9	BEGIN_MONTHS	2.952631
15	years_of_experience	2.679668
11	IS_MALE	2.282411
12	OWNS_CAR	2.188178
6	FLAG_PHONE	1.895831
5	FLAG_WORK_PHONE	1.785429
0	CNT_CHILDREN	1.610984
2	NAME_EDUCATION_TYPE	1.496645
3	MARITAL_STATUS	1.427603
4	HOUSE_TYPE	1.252956
7	FLAG_EMAIL	1.172928

Output 1.1. VIF Values obtained after the cleaned data

1.2. Assumption of no outliers

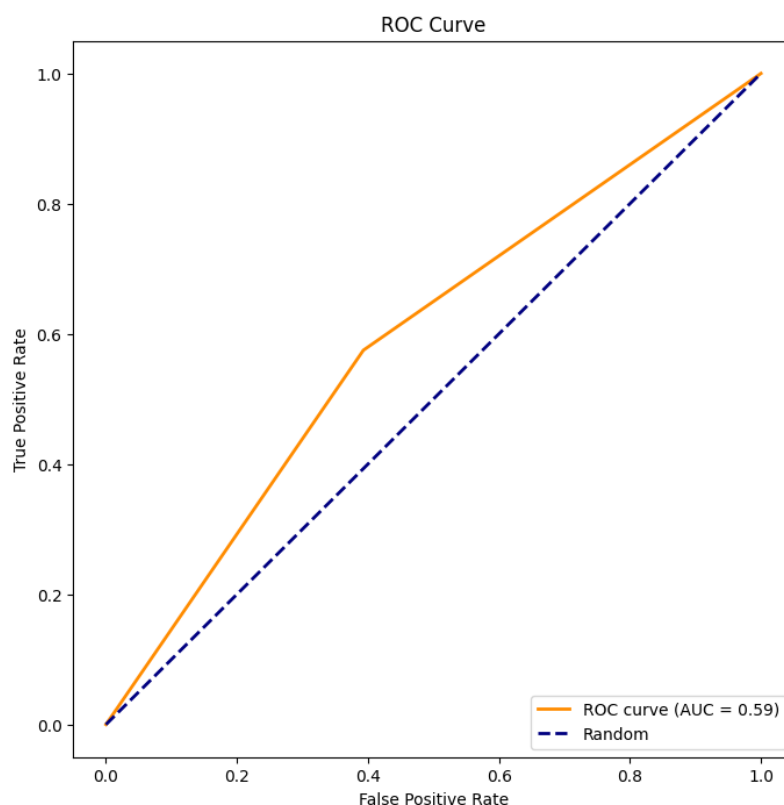
Outlier handling step I have done earlier where I removed the outliers in all the numerical columns and I presented the cleaned dataset as input here.

2. PERFORMANCE EVALUATION

2.1. CLASSIFICATION REPORT

Classification Report:					
		precision	recall	f1-score	support
	0	0.61	0.50	0.55	102414
	1	0.57	0.68	0.62	102472
accuracy				0.59	204886
macro avg		0.59	0.59	0.58	204886
weighted avg		0.59	0.59	0.58	204886

2.2. ROC-AUC PLOT



OUTPUT 2.2. ROC-AUC Curve for Test set

An AUC (Area Under the Receiver Operating Characteristic Curve) value of 0.59 is indicative of a model that has only slightly better-than-random performance in distinguishing between positive and negative classes. The AUC ranges from 0 to 1, where a value of 0.5 suggests random performance (no discrimination between classes), and a value of 1.0 indicates perfect discrimination.