

Approve a credit card analysis

**T5bootcamp data since project
Faris Auid Alharbi**



 **Slide** Members



CONTENTS

- 01** Project goals
- 02** Dataset
- 03** Process data
- 04** modeling



01 Project goals

Based on clients order records, who would be good to approve a credit card request? in this project i tray to be answer and help banks chooses between many clients requests.



02 dataset

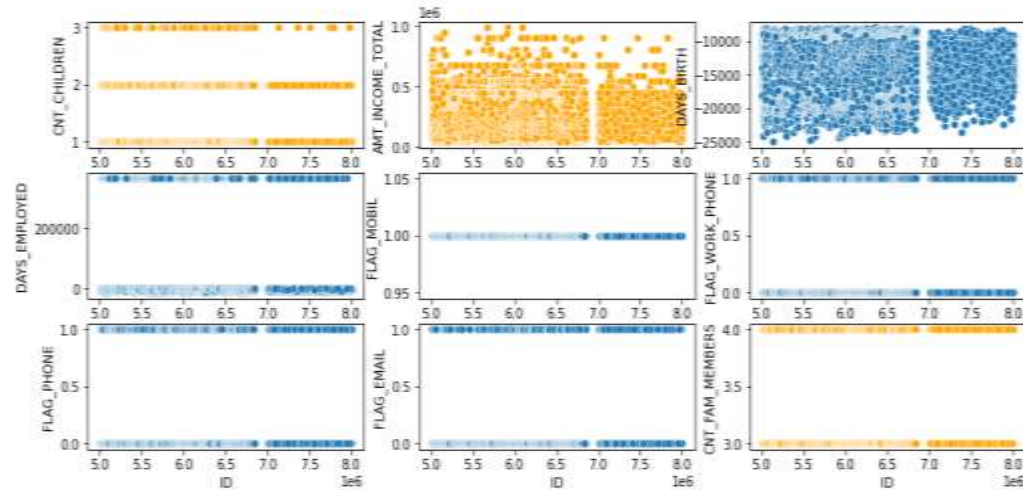
the Dataset is provided in .csv format from Kaggle it is 2 table .

1- application_record.csv ,it contains 438557 ID ,each ID has 18 Features .

2- credit_record.csv .it contains 1048575 ID ,each ID has 3 Features .

03 Process data

- Extracting data
- Clean data (null-non-numeric)
- Duplicate (ID)
- Remove features
- Combine between 2 table dataset
- most important feature (status)



04 modeling

A – XGBoost

is an optimized distributed gradient boosting library designed to be highly efficient, flexible and portable.

- Use because binary classification (metric)
- Accuracy : 93%
- Precision : 99%
- Recall: 87%
- F1: 93%

```
[47]: from sklearn.metrics import classification_report  
print(classification_report(y_test_balanced, prediction))
```

	precision	recall	f1-score	support
0	0.89	0.99	0.94	2805
1	0.99	0.87	0.93	2805
accuracy			0.93	5610
macro avg	0.94	0.93	0.93	5610
weighted avg	0.94	0.93	0.93	5610

04 modeling

B - (KNN) K-Nearest Neighbors Algorithm for Regression

- Accuracy : 68%
- Precision : 99%
- Recall: 36%
- F1: 53%

```
[49]: from sklearn.neighbors import KNeighborsClassifier  
knn1 = KNeighborsClassifier(n_neighbors=2)  
model = knn1.fit(X_balanced, y_balanced)  
prediction = knn1.predict(X_test_balanced)
```

```
[50]: print(classification_report(y_test_balanced, prediction))
```

	precision	recall	f1-score	support
0	0.61	0.99	0.76	2805
1	0.99	0.36	0.53	2805
accuracy			0.68	5610
macro avg	0.80	0.68	0.64	5610
weighted avg	0.80	0.68	0.64	5610



THANK YOU

CREDIT CARD



1234 5678 9876 5432

VALID
THRU 09/19