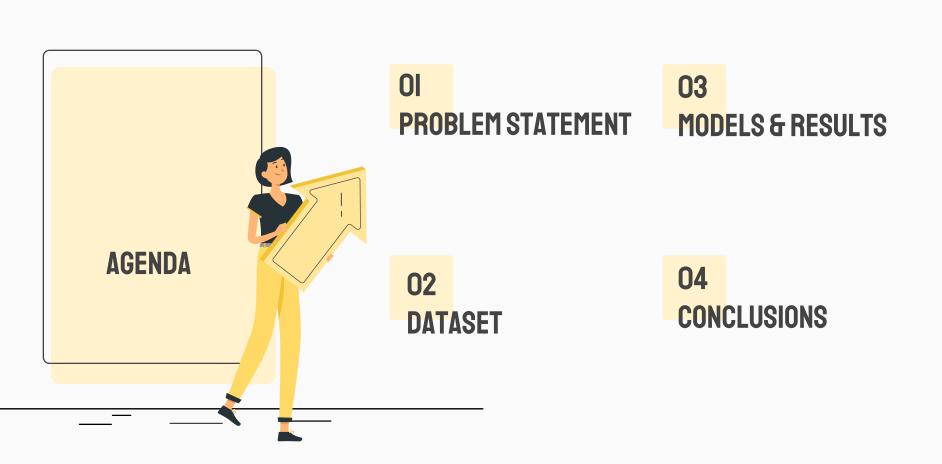
LOAN DEFAULTER PREDICTION

By: Maha Alatifi

13 January 2022





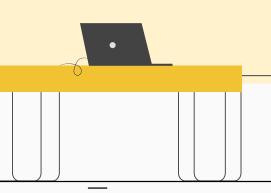




The main idea of this project was to develop a framework by using classification models to predict the loan defaulter in banking to minimise the risk of losing money while lending to customers.







OVERVIEW

- Data preparation and defining the correlation between features.
- > Training the dataset on 4 classifiers.
- Comparing between the models, in which the Random Forest showed the best score on F1.





307,511

APPLICATIONS 122

FEATURES

DATASET SUMMERY

SAMPLE DATA

1

0

0

0

171000.0

360000.0

112500.0

135000.0

1560726.0

1530000.0

1019610.0

405000.0

41301.0

42075.0

33826.5

20250.0

1395000.0

1530000.0

913500.0

405000.0

0	100002	1	Cash loans	М	N	Υ	0	202500.0	406597.5	24700.5	351000.0
1	100003	0	Cash loans	F	N	N	0	270000.0	1293502.5	35698.5	1129500.0
2	100004	0	Revolving loans	М	Υ	Υ	0	67500.0	135000.0	6750.0	135000.0
3	100006	0	Cash loans	F	N	Υ	0	135000.0	312682.5	29686.5	297000.0
4	100007	0	Cash loans	М	N	Υ	0	121500.0	513000.0	21865.5	513000.0
5	100008	0	Cash loans	М	N	Υ	0	99000.0	490495.5	27517.5	454500.0

Υ

Υ

Υ

100009

100010

100011

100012

6

7

8

9

0

0

0

0

Cash loans

Cash loans

Cash loans

Revolving loans

F

М

F

М

Υ

Υ

Ν

Ν

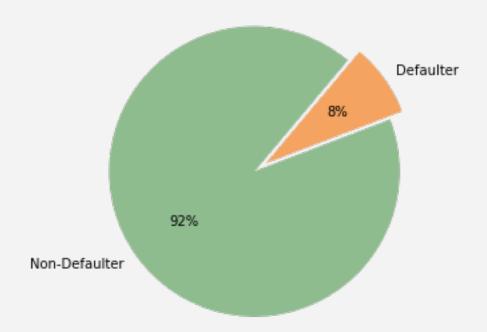
SK_ID_CURR TARGET NAME_CONTRACT_TYPE CODE_GENDER FLAG_OWN_CAR FLAG_OWN_REALTY CNT_CHILDREN AMT_INCOME_TOTAL AMT_CREDIT AMT_ANNUITY AMT_GOODS_PRICE

MODELS AND RESULTS



TARGET FEATURE DISTRIBUTION

Percentage of Target variable



FEATURES CORRELATIONS

																						 -10									
FLAG_DOCUMENT_2	1	-0.01	-5.9e-05	-0.00081	-0.002	-9e-05	-0.0019	-0.00041	-3.1e-05	-0.00041	-1.7e-05	-0.00039	-0.00035	-0.00023	-0.00065	-0.00011	-0.00059	-0.00016	-0.00015	-0.00012	0.0054										- 1.0
FLAG_DOCUMENT_3	-0.01	1	-0.014	-0.19	-0.49	-0.022	-0.47	-0.098	-0.0075	-0.093	-0.0012	-0.02	0.0021	0.00059	0.033	0.0025	0.008	0.0094	0.0074	-0.026	0.044		FLAG_MOBIL -	- 1	-0.00085	0.0009	-7.8e-05	0.0011	0.00044	0.00053	
FLAG_DOCUMENT_4	-5.9e-05	-0.014	1	-0.0011	-0.0028	-0.00012	-0.0027	-0.00056	-4.3e-05	-0.00057	-2.3e-05	-0.00054	-0.00049	-0.00031	-0.0009	-0.00015	-0.00082	-0.00022	-0.0002	-0.00017	-0.0027	- 0.8									
FLAG_DOCUMENT_5	-0.00081		-0.0011	1	-0.038	-0.0017	-0.037	-0.0077	-0.00059	-0.0078	-0.00032	-0.0074	-0.0067	-0.0043	-0.012	-0.002	-0.011	-0.003	-0.0028	0.00065	-0.00032		FLAG EMP PHONE -	-0 00085	1	0.23	-0.013	-0.016	0.063	0.046	- 0.8
FLAG_DOCUMENT_6	-0.002		-0.0028	-0.038	1	-0.0043	-0.092	-0.019	-0.0015	-0.019	-0.00079	-0.018	-0.014	-0.0098	-0.027	-0.0044	-0.024	-0.0048	-0.006	-0.0057	-0.029		TEAG_EINI_THORE	0.00005			0.025	0.020	0.005	0.010	
FLAG_DOCUMENT_7	-9e-05	-0.022	-0.00012	-0.0017	-0.0043	1	-0.0041	-0.00087	-6.6e-05	-0.00087	-3.5e-05	-0.00082	-0.00075	-0.00048	-0.0014	-0.00023	-0.0013	-0.00034	-0.00031	-0.00025	-0.0015	- 0.6									
FLAG_DOCUMENT_8	-0.0019	-0.47	-0.0027	-0.037	-0.092	-0.0041	1	-0.019	-0.0014	-0.018	0.0039	0.077	0.03	0.024	0.013	0.0053	0.0078	5.3e-05	0.0039	-0.0015	-0.008		FLAG_WORK_PHONE -	0.0009			0.022	0.29	-0.012	0.029	- 0.6
FLAG_DOCUMENT_9	-0.00041	-0.098	-0.00056	-0.0077	-0.019	-0.00087	-0.019	1	-0.0003	-0.0039	-0.00016	0.00068	0.0014	0.0023	0.0085	0.0086	0.0019	-0.0015	-0.0014	-0.0011	-0.0044										- 0.0
FLAG_DOCUMENT_10	-3.1e-05	-0.0075	4.3e-05	-0.00059	-0.0015	-6.6e-05	-0.0014	-0.0003	1	-0.0003	-1.2e-05	-0.00028	-0.00026	-0.00017	-0.00048	-7.8e-05	-0.00043	-0.00012	-0.00011	-8.7e-05	-0.0014	- 0.4	FLAG CONT MOBILE -	-7 8e-05	A) () 13	0.022	,	0.0063	-0.0054	0.00037	
FLAG_DOCUMENT_11	-0.00041	-0.093	-0.00057	-0.0078	-0.019	-0.00087	-0.018	-0.0039	-0.0003	1	-0.00016	-0.0037	-0.0034	-0.0022	-0.0063	-0.001	-0.0057	-0.0015	-0.0014	-0.0011	-0.0042		TENG_CONT_MODILE	-7.00-03	-0.015	0.022	-	0.0003	-0.0054	0.00037	
FLAG_DOCUMENT_12	-1.7e-05	-0.0012	-2.3e-05	-0.00032	-0.00079	-3.5e-05	0.0039	-0.00016	-1.2e-05	-0.00016	1	-0.00015	-0.00014	-8.9e-05	-0.00026	-4.2e-05	-0.00023	-6.2e-05	-5.7e-05	-4.7e-05	-0.00076										- 0.4
FLAG_DOCUMENT_13	-0.00039	-0.02	-0.00054	-0.0074	-0.018	-0.00082	0.077	0.00068	-0.00028	-0.0037	-0.00015	1	-0.0032	-0.0021	-0.006	-0.00097	-0.0054	-0.0015	0.043	0.0049	-0.012	- 0.2	FLAG_PHONE -	0.0011	-0.016	0.29	0.0063	1	0.015	-0.024	
FLAG_DOCUMENT_14	-0.00035	0.0021	-0.00049	-0.0067	-0.014	-0.00075	0.03	0.0014	-0.00026	-0.0034	-0.00014	-0.0032		-0.0019	-0.0054	-0.00089	-0.0049	-0.0013	0.031	0.0023	-0.0095										
FLAG_DOCUMENT_15	-0.00023	0.00059	-0.00031	-0.0043	-0.0098	-0.00048	0.024	0.0023	-0.00017	-0.0022	-8.9e-05	-0.0021	-0.0019	1	-0.0035	-0.00057	-0.0032	-0.00085	0.028	-0.00064	-0.0065		FLAG EMAIL -	0.00044	0.063	-0.012	-0.0054	0.015	1	-0.0018	- 0.2
FLAG_DOCUMENT_16	-0.00065	0.033	-0.0009	-0.012	-0.027	-0.0014	0.013	0.0085	-0.00048	-0.0063	-0.00026	-0.006	-0.0054	-0.0035	1	-0.0016	-0.0091	-0.0024	0.082	0.0035	-0.012	-0.0	FLAG_EMAIL -	0.00044	0.063	-0.012	-0.0054	0.015	1	-0.0016	- 0.2
FLAG_DOCUMENT_17	-0.00011	0.0025	-0.00015	-0.002	-0.0044	-0.00023	0.0053	0.0086	-7.8e-05	-0.001	-4.2e-05	-0.00097	-0.00089	-0.00057	-0.0016	1	-0.0015	-0.0004	0.026	-0.0003	-0.0034										
FLAG_DOCUMENT_18	-0.00059	0.008	-0.00082	-0.011	-0.024	-0.0013	0.0078	0.0019	-0.00043	-0.0057	-0.00023	-0.0054	-0.0049	-0.0032	-0.0091	-0.0015	1	-0.0022	0.082	0.0063	-0.008	0.2	TARGET -	0.00053	0.046	0.029	0.00037	-0.024	-0.0018	1	
FLAG_DOCUMENT_19	-0.00016	0.0094	-0.00022	-0.003	-0.0048	-0.00034	5.3e-05	-0.0015	-0.00012	-0.0015	-6.2e-05	-0.0015	-0.0013	-0.00085	-0.0024	-0.0004	-0.0022	1	0.035	-0.00045	-0.0014	0.2		لحبط			لسبسا				- 0.0
FLAG_DOCUMENT_20	-0.00015	0.0074	-0.0002	-0.0028	-0.006	-0.00031	0.0039	-0.0014	-0.00011	-0.0014	-5.7e-05	0.043	0.031	0.028	0.082	0.026	0.082	0.035	1	0.0075	0.00022			BIL.	PHONE	PHONE	SIE.	PHONE	IAIL.	TARGET .	
FLAG_DOCUMENT_21	-0.00012	-0.026	-0.00017	0.00065	-0.0057	-0.00025	-0.0015	-0.0011	-8.7e-05	-0.0011	-4.7e-05	0.0049	0.0023	-0.00064	0.0035	-0.0003	0.0063	-0.00045	0.0075	1	0.0037	0.4		AG_MOBIL	H.	Ę.	CONT_MOBIL	H.	AG_EMAIL	TAR	
TARGET -	0.0054	0.044	-0.0027	-0.00032	-0.029	-0.0015	-0.008	-0.0044	-0.0014	-0.0042	-0.00076	-0.012	-0.0095	-0.0065	-0.012	-0.0034	-0.008	-0.0014	0.00022	0.0037	1	0.4		FLA	EMP	VORK	TNO	FLAG_	FF		
	T_2-	ė	- 4 <u>-</u>	-5_	- 9 ⁻ L	-7.	8	-6_T	- 00	Ė,	17	ė,	14	-51	16 -	17 -	18	-19	- 20 -	- 21 -	- EE				ILAG	4G_V					
	CUMEN	CUMEN	CUMEN	CUMEN	CUMEN	CUMEN	CUMEN	CUMEN	UMENT	UMENT	UMENT	UMENT	UMENT	TARC				ш	5	5											
	LAG_DOI	LAG_DOI	'LAG_DOI	LAG_DOI	'LAG_DOI	LAG_DOI	LAG_DOI	'LAG_DOI	AG_DOC	AG_DOO	AG_DOC	AG_DOC	AG_DOC	AG_DOC	AG_DOC	AG_DOC	AG_DOO	AG_DOC	AG_DOC	AG_DOC											
TARGET -	FLAG_DOCUMENT_2 - 600	FLAG_DOCUMENT_3 - 60	FLAG_DOCUMENT_4 -	FLAG_DOCUMENT_S -	FLAG_DOCUMENT_6 -	FLAG_DOCUMENT_7 -	FLAG_DOCUMENT_8 - 80	FLAG_DOCUMENT_9 -	FLAG_DOCUMENT_10 -	FLAG_DOCUMENT_11 -	FLAG_DOCUMENT_12 -	FLAG_DOCUMENT_13 -	FLAG_DOCUMENT_14 -	FLAG_DOCUMENT_15 -	FLAG_DOCUMENT_16 -	FLAG_DOCUMENT_17 - PEOP	FLAG_DOCUMENT_18 -	FLAG_DOCUMENT_19 -	FLAG_DOCUMENT_20 -	FLAG_DOCUMENT_21 -	TARGET			료	FLAG_E	FLAG_WORK_	FLAG_COI	교	Ľ.		

THE CLASSIFIERS MODELS



- ✓ Logistic regression (LR)
- ✓ K-nearest neighbours (K-NN)
- ✓ Decision Tree (DT)
- ✓ Random Forest (RF)





Confusion Matrix

Predicted Values

Positive

Negative

Positive

TP

Actual: Defaulter Predict: Defaulter

FP

Actual: Non_Defaulter Predict: Defaulter

Negative

FN

Actual: Defaulter Predict: Non_Defaulter

ΤN

Actual: Non_Defaulter Predict: Non_Defaulter

RESULTS FOR 4 MODELS

Model	Function	Precision	Recall	Accuracy	F1
Logistic regression (LR)	LogisticRegression	0.87	0.54	0.91	0.64
K-nearest neighbours (K-NN)	KNeighborsClassifier	0.86	0.91	0.92	0.84
Decision Tree (DT)	DecisionTreeClassifier	0.88	0.61	0.84	0.70
Random Forest (RF)	RandomForestClassifier	0.85	0.92	0.92	0.88

THANKS!

Do you have any questions?

