





CREDIT RISK PREDICTION

BY M. RIFKI OSKAR

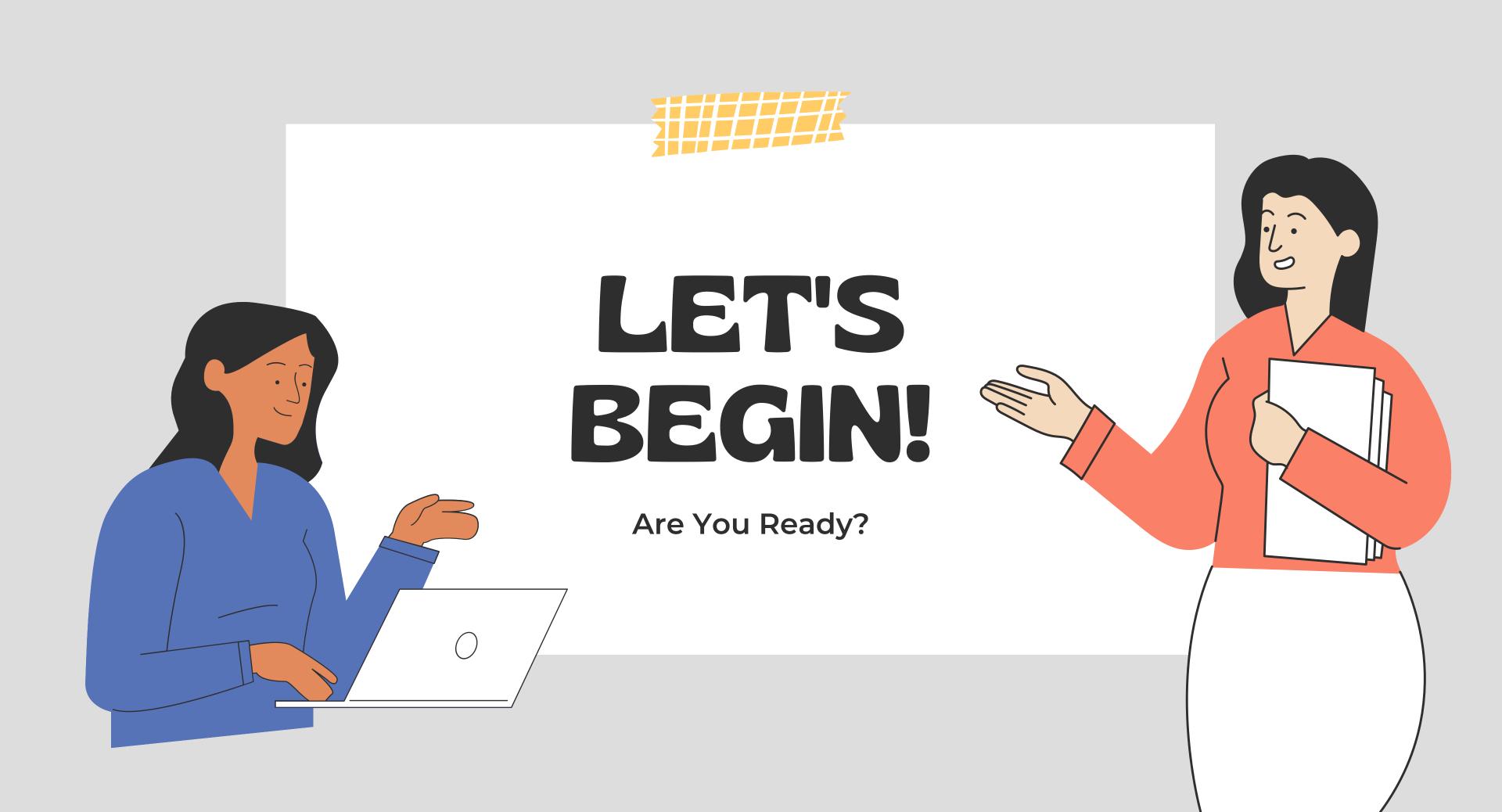


1 Background

2 Exploratory Data Analysis

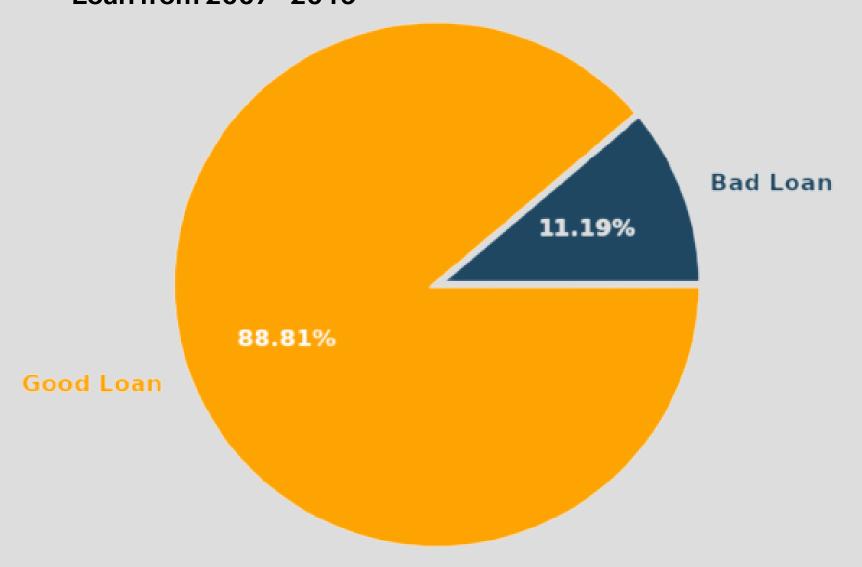
B Data Preprocessing

4 Modelling



PROBLEM STATEMENT

We get 88.81% Good Loan and 11.19% Bad Loan from 2007 - 2016



BACKGROUND

GOAL

Reducing the number of customers with bad loan (Bad Loan)

OBJECTIVE

Create machine learning models to conduct profiling or predict customer profiles that have the potential to bad loans so that companies are able to carry out more accurate customer assessments and improve company work performance

BUSINESS METRIC



Default Loans (Low or High)

EXPLORATORY DATA ANALYSIS

Loan data

75 Number of Features

466285 Number of Rows

Numerical Features: 52

0Unnamed: 0466285 non-null int641id466285 non-null int642member_id466285 non-null int643loan_amnt466285 non-null int644funded_amnt466285 non-null int645funded_amnt_inv466285 non-null float646int_rate466285 non-null float647installment466285 non-null float648annual_inc466281 non-null float649dti466285 non-null float6410delinq_2yrs466256 non-null float6411inq_last_6mths466256 non-null float6412mths_since_last_delinq215934 non-null float6413mths_since_last_record62638 non-null float6414open_acc466256 non-null float6415pub_rec466256 non-null float6416revol_bal466285 non-null float6417revol_util465945 non-null float6418total_acc466285 non-null float6419out_prncp466285 non-null float6420out_prncp_inv466285 non-null float6421total_pymnt_inv466285 non-null float6422total_rec_prncp466285 non-null float6424total_rec_late_fee466285 non-null float6425total_rec_late_fee466285 non-null float6426recoveries466285 non-null float6427collection_recovery_fee466285 non-null float6428last_pymnt_amnt466285 non-null float6429				
2 member_id 3 loan_amnt 466285 non-null int64 4 funded_amnt 5 funded_amnt 466285 non-null int64 6 int_rate 466285 non-null float64 7 installment 8 annual_inc 9 dti 10 delinq_2yrs 11 delinq_2yrs 12 mths_since_last_delinq 13 mths_since_last_record 16 revol_bal 17 revol_util 18 total_acc 19 out_prncp 20 out_prncp_inv 21 total_pymnt 21 total_pymnt_inv 466285 non-null float64	0	Unnamed: 0	466285 non-null	int64
loan_amnt	1	id	466285 non-null	int64
4 funded_amnt	2	member_id	466285 non-null	int64
5 funded_amnt_inv 466285 non-null float64 6 int_rate 466285 non-null float64 7 installment 466285 non-null float64 8 annual_inc 466281 non-null float64 9 dti 466285 non-null float64 10 delinq_2yrs 466256 non-null float64 11 inq_last_6mths 466256 non-null float64 12 mths_since_last_record 62638 non-null float64 13 mths_since_last_record 62638 non-null float64 14 open_acc 466256 non-null float64 15 pub_rec 466256 non-null float64 16 revol_bal 466285 non-null float64 17 revol_util 465945 non-null float64 18 total_acc 466256 non-null float64 19 out_prncp 466285 non-null float64 20 out_prncp_inv 466285 non-null float64 21 total_pymnt 466285 non-null float64 22 total_pymnt_inv 466285 non-null float64 23 total_rec_prncp 466285 non-null float64 24 total_rec_int 466285 non-null float64 25 total_rec_late_fee 466285 non-null float64 26 recoveries 466285 non-null float64 27 collection_recovery_fee 466285 non-null float64 29 collection_s12_mths_ex_med 466285 non-null float64 30 mths_since_last_major_derog 88974 non-null float64 30 mths_since_last_major_derog 88974 non-null float64 31 policy_code 466285 non-null float64 32 annual_inc_joint 9 non-null float64 33 dti_joint 9 non-null float64 34 verification_status_joint 9 non-null float64 35 acc_now_deling 466256 non-null float64 37 tot_cur_bal 396009 non-null float64 37 tot_cur_bal	3	loan_amnt	466285 non-null	int64
6 int_rate 7 installment 8 annual_inc 9 dti 10 delinq_2yrs 11 inq_last_6mths 12 mths_since_last_record 15 pub_rec 16 revol_bal 17 total_acc 18 total_acc 19 out_prncp 20 out_prncp_inv 21 total_pymnt 21 total_pymnt 21 total_pymnt_inv 22 total_rec_int 23 total_rec_last_ment 24 total_rec_int 25 collections_12_mths_ex_med 36 mths_since_last_ment 37 tot_cur_bal 466285 non-null 4662864 466285 non-null 466286 466285 non-null	4	funded_amnt	466285 non-null	int64
7 installment 466285 non-null float64 8 annual_inc 466281 non-null float64 9 dti 466285 non-null float64 10 delinq_2yrs 466256 non-null float64 11 inq_last_6mths 466256 non-null float64 12 mths_since_last_delinq 215934 non-null float64 13 mths_since_last_record 62638 non-null float64 14 open_acc 466256 non-null float64 15 pub_rec 466256 non-null float64 16 revol_bal 466285 non-null float64 17 revol_util 465945 non-null float64 18 total_acc 466256 non-null float64 19 out_prncp 466285 non-null float64 20 out_prncp 466285 non-null float64 21 total_pymnt 466285 non-null float64 22 total_pymnt_inv 466285 non-null float64 23 total_rec_prncp 466285 non-null float64 24 total_rec_int 466285 non-null float64 25 total_rec_late_fee 466285 non-null float64 26 recoveries 466285 non-null float64 27 collection_recovery_fee 466285 non-null float64 28 last_pymnt_amnt 466285 non-null float64 29 collections_12_mths_ex_med 466140 non-null float64 30 mths_since_last_major_derog 98974 non-null float64 31 policy_code 466285 non-null float64 32 annual_inc_joint 0 non-null float64 33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	5	<pre>funded_amnt_inv</pre>	466285 non-null	float64
8 annual_inc 9 dti 466285 non-null float64 10 delinq_2yrs 466256 non-null inq_last_6mths 466256 non-null intoat64 11 inq_last_femths 466256 non-null float64 12 mths_since_last_delinq 13 mths_since_last_record 62638 non-null float64 14 open_acc 466256 non-null float64 15 pub_rec 466256 non-null float64 16 revol_bal 466285 non-null float64 17 revol_util 465945 non-null float64 18 total_acc 466256 non-null float64 19 out_prncp 466285 non-null float64 20 out_prncp_inv 466285 non-null float64 21 total_pymnt 466285 non-null float64 22 total_pymnt 466285 non-null float64 23 total_rec_prncp 466285 non-null float64 24 total_rec_int 466285 non-null float64 25 total_rec_late_fee 466285 non-null float64 26 recoveries 466285 non-null float64 27 collection_recovery_fee 466285 non-null float64 28 last_pymnt_amnt 466285 non-null float64 29 collection_recovery_fee 466285 non-null float64 30 mths_since_last_major_derog 31 policy_code 466285 non-null float64 32 annual_inc_joint 0 non-null float64 33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal	6	int_rate	466285 non-null	float64
9 dti 466285 non-null float64 10 delinq_2yrs 466256 non-null float64 11 inq_last_6mths 466256 non-null float64 12 mths_since_last_delinq 215934 non-null float64 13 mths_since_last_record 62638 non-null float64 14 open_acc 466256 non-null float64 15 pub_rec 466256 non-null float64 16 revol_bal 466285 non-null int64 17 revol_util 465945 non-null float64 18 total_acc 466256 non-null float64 19 out_prncp 466285 non-null float64 20 out_prncp_inv 466285 non-null float64 21 total_pymnt 466285 non-null float64 22 total_pymnt_inv 466285 non-null float64 23 total_rec_prncp 466285 non-null float64 24 total_rec_int 466285 non-null float64 25 total_rec_late_fee 466285 non-null float64 26 recoveries 466285 non-null float64 27 collection_recovery_fee 466285 non-null float64 28 last_pymnt_amnt 466285 non-null float64 29 collection_recovery_fee 466285 non-null float64 30 mths_since_last_major_derog 98974 non-null float64 31 policy_code 466285 non-null float64 32 annual_inc_joint 0 non-null float64 33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	7	installment	466285 non-null	float64
delinq_2yrs	8	annual_inc	466281 non-null	float64
11inq_last_6mths466256 non-nullfloat6412mths_since_last_delinq215934 non-nullfloat6413mths_since_last_record62638 non-nullfloat6414open_acc466256 non-nullfloat6415pub_rec466256 non-nullfloat6416revol_bal466285 non-nullint6417revol_util465945 non-nullfloat6418total_acc466256 non-nullfloat6419out_prncp466285 non-nullfloat6420out_prncp_inv466285 non-nullfloat6421total_pymnt466285 non-nullfloat6422total_pymnt_inv466285 non-nullfloat6423total_rec_prncp466285 non-nullfloat6424total_rec_late_fee466285 non-nullfloat6425total_rec_late_fee466285 non-nullfloat6426recoveries466285 non-nullfloat6427collection_recovery_fee466285 non-nullfloat6428last_pymnt_amnt466285 non-nullfloat6429collections_12_mths_ex_med466140 non-nullfloat6430mths_since_last_major_derog98974 non-nullfloat6431policy_code466285 non-nullfloat6432annual_inc_joint0 non-nullfloat6433dti_joint0 non-nullfloat6434verification_status_joint0 non-nullfloat6436tot_coll_amt <td< td=""><td>9</td><td>dti</td><td>466285 non-null</td><td>float64</td></td<>	9	dti	466285 non-null	float64
11 inq_last_6mths 466256 non-null float64 12 mths_since_last_delinq 215934 non-null float64 13 mths_since_last_record 62638 non-null float64 14 open_acc 466256 non-null float64 15 pub_rec 466256 non-null float64 16 revol_bal 466285 non-null int64 17 revol_util 465945 non-null float64 18 total_acc 466285 non-null float64 19 out_prncp 466285 non-null float64 20 out_prncp_inv 466285 non-null float64 21 total_pymnt_inv 466285 non-null float64 22 total_pymnt_inv 466285 non-null float64 23 total_rec_prncp 466285 non-null float64 24 total_rec_late_fee 466285 non-null float64 25 total_rec_late_fee 466285 non-null float64 26 recoveries 466285 non-null float64 27 collections_12_mths_ex_med 466140 non-null f	10	delinq_2yrs	466256 non-null	float64
13 mths_since_last_record 14 open_acc 15 pub_rec 16 revol_bal 17 revol_util 18 total_acc 19 out_prncp 20 ut_prncp 21 total_pymnt 22 total_pymnt_inv 23 total_rec_int 24 total_rec_int 25 total_rec_late_fee 26 recoveries 27 collection_recovery_fee 28 last_pymnt_amnt 29 collections_12_mths_ex_med 30 mths_since_last_major_derog 31 policy_code 32 annual_inc_joint 33 e6029 non-null float64 34 tot_coll_amt 35 acc_now_delinq 466285 non-null float64	11		466256 non-null	float64
14 open_acc 466256 non-null float64 15 pub_rec 466256 non-null float64 16 revol_bal 466285 non-null int64 17 revol_util 465945 non-null float64 18 total_acc 466256 non-null float64 19 out_prncp 466285 non-null float64 20 out_prncp_inv 466285 non-null float64 21 total_pymnt_inv 466285 non-null float64 22 total_pymnt_inv 466285 non-null float64 23 total_rec_prncp 466285 non-null float64 24 total_rec_int 466285 non-null float64 25 total_rec_late_fee 466285 non-null float64 26 recoveries 466285 non-null float64 27 collection_recovery_fee 466285 non-null float64 28 last_pymnt_amnt 466285 non-null float64 29 collections_12_mths_ex_med 466140 non-null float64 30 mths_since_last_major_derog 98974 non-null float64 31 policy_code 466285 non-null float64 32 annual_inc_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36009 non-null float64 401_cate 37 tot_cur_bal	12	mths_since_last_delinq	215934 non-null	float64
pub_rec	13	mths_since_last_record	62638 non-null	float64
16 revol_bal 466285 non-null int64 17 revol_util 465945 non-null float64 18 total_acc 466256 non-null float64 19 out_prncp 466285 non-null float64 20 out_prncp_inv 466285 non-null float64 21 total_pymnt 466285 non-null float64 22 total_pymnt_inv 466285 non-null float64 23 total_rec_prncp 466285 non-null float64 24 total_rec_int 466285 non-null float64 25 total_rec_late_fee 466285 non-null float64 26 recoveries 466285 non-null float64 27 collection_recovery_fee 466285 non-null float64 28 last_pymnt_amnt 466285 non-null float64 29 collections_12_mths_ex_med 466140 non-null float64 30 mths_since_last_major_derog 98974 non-null float64 31 policy_code 466285 non-null float64 32 annual_inc_joint 0 non-null float64 33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal	14	open_acc	466256 non-null	float64
17 revol_util 465945 non-null float64 18 total_acc 466256 non-null float64 19 out_prncp 466285 non-null float64 20 out_prncp_inv 466285 non-null float64 21 total_pymnt 466285 non-null float64 22 total_pymnt_inv 466285 non-null float64 23 total_rec_prncp 466285 non-null float64 24 total_rec_int 466285 non-null float64 25 total_rec_late_fee 466285 non-null float64 26 recoveries 466285 non-null float64 27 collection_recovery_fee 466285 non-null float64 28 last_pymnt_amnt 466285 non-null float64 29 collections_12_mths_ex_med 466140 non-null float64 30 mths_since_last_major_derog 98974 non-null float64 31 policy_code 466285 non-null float64 32 annual_inc_joint 0 non-null float64 33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	15	pub_rec	466256 non-null	float64
18total_acc466256 non-nullfloat6419out_prncp466285 non-nullfloat6420out_prncp_inv466285 non-nullfloat6421total_pymnt466285 non-nullfloat6422total_pymnt_inv466285 non-nullfloat6423total_rec_prncp466285 non-nullfloat6424total_rec_int466285 non-nullfloat6425total_rec_late_fee466285 non-nullfloat6426recoveries466285 non-nullfloat6427collection_recovery_fee466285 non-nullfloat6428last_pymnt_amnt466285 non-nullfloat6429collections_12_mths_ex_med466140 non-nullfloat6430mths_since_last_major_derog98974 non-nullfloat6431policy_code466285 non-nullint6432annual_inc_joint0 non-nullfloat6433dti_joint0 non-nullfloat6434verification_status_joint0 non-nullfloat6435acc_now_delinq466256 non-nullfloat6436tot_coll_amt396009 non-nullfloat6437tot_cur_bal396009 non-nullfloat64	16	revol_bal	466285 non-null	int64
19 out_prncp	17	revol_util	465945 non-null	float64
20out_prncp_inv466285 non-nullfloat6421total_pymnt466285 non-nullfloat6422total_pymnt_inv466285 non-nullfloat6423total_rec_prncp466285 non-nullfloat6424total_rec_int466285 non-nullfloat6425total_rec_late_fee466285 non-nullfloat6426recoveries466285 non-nullfloat6427collection_recovery_fee466285 non-nullfloat6428last_pymnt_amnt466285 non-nullfloat6429collections_12_mths_ex_med466140 non-nullfloat6430mths_since_last_major_derog98974 non-nullfloat6431policy_code466285 non-nullfloat6432annual_inc_joint0 non-nullfloat6433dti_joint0 non-nullfloat6434verification_status_joint0 non-nullfloat6435acc_now_delinq466256 non-nullfloat6436tot_coll_amt396009 non-nullfloat6437tot_cur_bal396009 non-nullfloat64	18	total_acc	466256 non-null	float64
total_pymnt	19	out_prncp	466285 non-null	float64
total_pymnt_inv 466285 non-null float64 total_rec_prncp 466285 non-null float64 total_rec_int 466285 non-null float64 total_rec_late_fee 466285 non-null float64 collection_recovery_fee 466285 non-null float64 total_pymnt_amnt 466285 non-null float64 collections_12_mths_ex_med 466140 non-null float64 mths_since_last_major_derog 98974 non-null float64 mths_since_last_major_derog 98974 non-null float64 annual_inc_joint 0 non-null float64 dti_joint 0 non-null float64 dti_joint 0 non-null float64 dti_joint 0 non-null float64 dti_joint 0 non-null float64 dti_coll_amt 396009 non-null float64 tot_coll_amt 396009 non-null float64 dti_cur_bal 396009 non-null float64	20	out_prncp_inv	466285 non-null	float64
total_rec_prncp 466285 non-null float64 total_rec_int 466285 non-null float64 total_rec_late_fee 466285 non-null float64 recoveries 466285 non-null float64 recoveries 466285 non-null float64 recoveries 466285 non-null float64 last_pymnt_amnt 466285 non-null float64 recoveries 466285 non-null flo	21	total_pymnt	466285 non-null	float64
24total_rec_int466285 non-null float6425total_rec_late_fee466285 non-null float6426recoveries466285 non-null float6427collection_recovery_fee466285 non-null float6428last_pymnt_amnt466285 non-null float6429collections_12_mths_ex_med466140 non-null float6430mths_since_last_major_derog98974 non-null float6431policy_code466285 non-null float6432annual_inc_joint0 non-null float6433dti_joint0 non-null float6434verification_status_joint0 non-null float6435acc_now_delinq466256 non-null float6436tot_coll_amt396009 non-null float6437tot_cur_bal396009 non-null float64	22	total_pymnt_inv	466285 non-null	float64
total_rec_late_fee 466285 non-null float64 recoveries 466285 non-null float64 recoveries 466285 non-null float64 recoveries 466285 non-null float64 last_pymnt_amnt 466285 non-null float64 recoveries 466285 non-null float64 rec	23	total_rec_prncp	466285 non-null	float64
26 recoveries 466285 non-null float64 27 collection_recovery_fee 466285 non-null float64 28 last_pymnt_amnt 466285 non-null float64 29 collections_12_mths_ex_med 466140 non-null float64 30 mths_since_last_major_derog 98974 non-null float64 31 policy_code 466285 non-null int64 32 annual_inc_joint 0 non-null float64 33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	24	total_rec_int	466285 non-null	float64
27 collection_recovery_fee 28 last_pymnt_amnt 466285 non-null float64 29 collections_12_mths_ex_med 30 mths_since_last_major_derog 31 policy_code 32 annual_inc_joint 33 dti_joint 466285 non-null float64 36 tot_coll_amt 37 tot_cur_bal 466285 non-null float64	25	total_rec_late_fee	466285 non-null	float64
28 last_pymnt_amnt 466285 non-null float64 29 collections_12_mths_ex_med 466140 non-null float64 30 mths_since_last_major_derog 98974 non-null float64 31 policy_code 466285 non-null int64 32 annual_inc_joint 0 non-null float64 33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	26	recoveries	466285 non-null	float64
29 collections_12_mths_ex_med 466140 non-null float64 30 mths_since_last_major_derog 98974 non-null float64 31 policy_code 466285 non-null int64 32 annual_inc_joint 0 non-null float64 33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	27	collection_recovery_fee	466285 non-null	float64
30mths_since_last_major_derog98974 non-nullfloat6431policy_code466285 non-nullint6432annual_inc_joint0 non-nullfloat6433dti_joint0 non-nullfloat6434verification_status_joint0 non-nullfloat6435acc_now_delinq466256 non-nullfloat6436tot_coll_amt396009 non-nullfloat6437tot_cur_bal396009 non-nullfloat64	28	last_pymnt_amnt	466285 non-null	float64
31 policy_code 466285 non-null int64 32 annual_inc_joint 0 non-null float64 33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	29	collections_12_mths_ex_med	466140 non-null	float64
32 annual_inc_joint 0 non-null float64 33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	30	<pre>mths_since_last_major_derog</pre>	98974 non-null	float64
33 dti_joint 0 non-null float64 34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	31	policy_code	466285 non-null	int64
34 verification_status_joint 0 non-null float64 35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	32	annual_inc_joint	0 non-null	float64
35 acc_now_delinq 466256 non-null float64 36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	33	dti_joint	0 non-null	float64
36 tot_coll_amt 396009 non-null float64 37 tot_cur_bal 396009 non-null float64	34	verification_status_joint	0 non-null	float64
37 tot_cur_bal 396009 non-null float64	35	acc_now_delinq	466256 non-null	float64
	36	tot_coll_amt	396009 non-null	float64
38 open_acc_6m 0 non-null float64	37	tot_cur_bal	396009 non-null	float64
	38	open_acc_6m	0 non-null	float64

```
39 open_il_6m
                                0 non-null
                                                 float64
40 open_il_12m
                                0 non-null
                                                 float64
41 open_il_24m
                                0 non-null
                                                 float64
42 mths_since_rcnt_il
                                0 non-null
                                                 float64
43 total_bal_il
                                0 non-null
                                                 float64
44 il_util
                                0 non-null
                                                 float64
45 open_rv_12m
                                0 non-null
                                                 float64
                                                 float64
46 open_rv_24m
                                0 non-null
47 max_bal_bc
                                0 non-null
                                                 float64
                                                 float64
   all_util
                                0 non-null
49 total_rev_hi_lim
                                396009 non-null float64
50 inq_fi
                                0 non-null
                                                 float64
51 total_cu_tl
                                                 float64
                                0 non-null
52 inq_last_12m
                                0 non-null
                                                 float64
```

Target:
Loan Status -> Flag

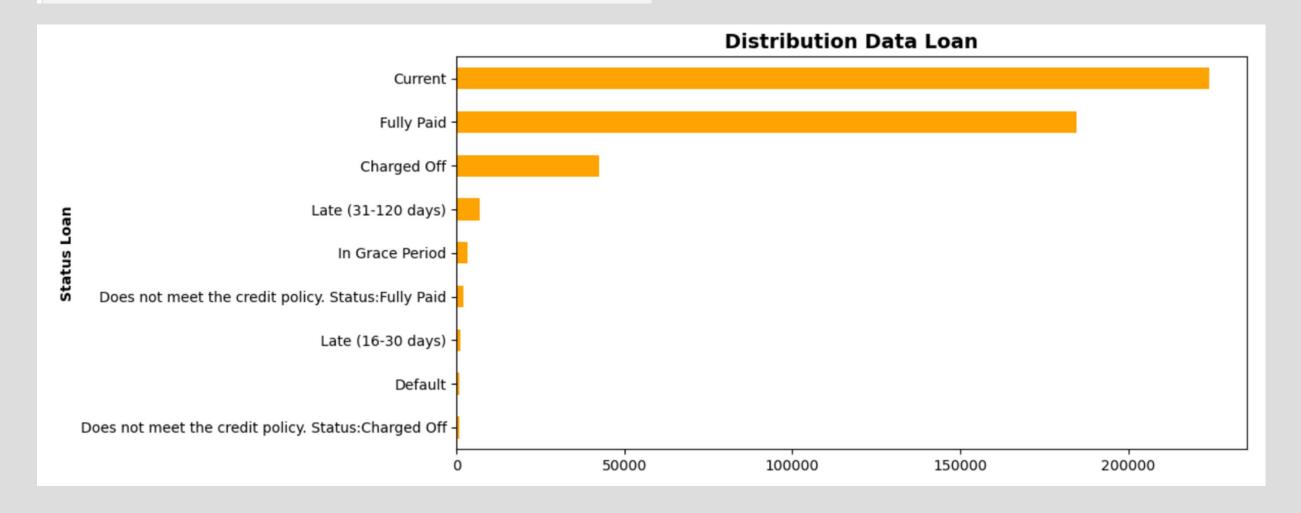
Categorical Features: 22

0	term	466285	non-null	object
1	grade	466285	non-null	object
2	sub_grade	466285	non-null	object
3	emp_title	438697	non-null	object
4	emp_length	445277	non-null	object
5	home_ownership	466285	non-null	object
6	verification_status	466285	non-null	object
7	issue_d	466285	non-null	object
8	loan_status	466285	non-null	object
9	pymnt_plan	466285	non-null	object
10	url	466285	non-null	object
11	desc	125983	non-null	object
12	purpose	466285	non-null	object
13	title	466265	non-null	object
14	zip_code	466285	non-null	object
15	addr_state	466285	non-null	object
16	earliest_cr_line	466256	non-null	object
17	initial_list_status	466285	non-null	object
18	last_pymnt_d	465909	non-null	object
19	next_pymnt_d	239071	non-null	object
20	<pre>last_credit_pull_d</pre>	466243	non-null	object
21	application_type	466285	non-null	object

BUSINESS INSIGHTS

	target	total	%
1	Current	224226	48.090000
5	Fully Paid	184739	39.620000
0	Charged Off	42475	9.110000
8	Late (31-120 days)	6900	1.480000
6	In Grace Period	3146	0.670000
4	Does not meet the credit policy. Status:Fully	1988	0.430000
7	Late (16-30 days)	1218	0.260000
2	Default	832	0.180000
3	Does not meet the credit policy. Status:Charge	761	0.160000

There are about 48% applicants with loan status current, followed by loan status fully paid 39.6%



PREPROCESSING

FEATURE ENGINEERING



Feature extraction

 Membuat feature baru dari feature yang sudah ada seperti feature Flag dimana akan menjadi feature target



Handling Missing Value

- Drop null values
- Handling error values



Handling duplicated data

• tidak terdapat data yang duplicate

PREPROCESSING



Feature selection

- Membuang feature-feature yang tidak relevan
- Membuang feature-feature yang redundan

DATA CLEANSING



Feature transformation

• melakukan standarization



Feature encoding

- mengubah feature flag menjadi numeric, 0 dan 1 secara manual
- mengubah feature categorical menjadi numeric menggunakan one hot encoding



Handle class imbalance

 Dataset tidak dilakukan handling karena imbalance tidak terlalu ekstrem

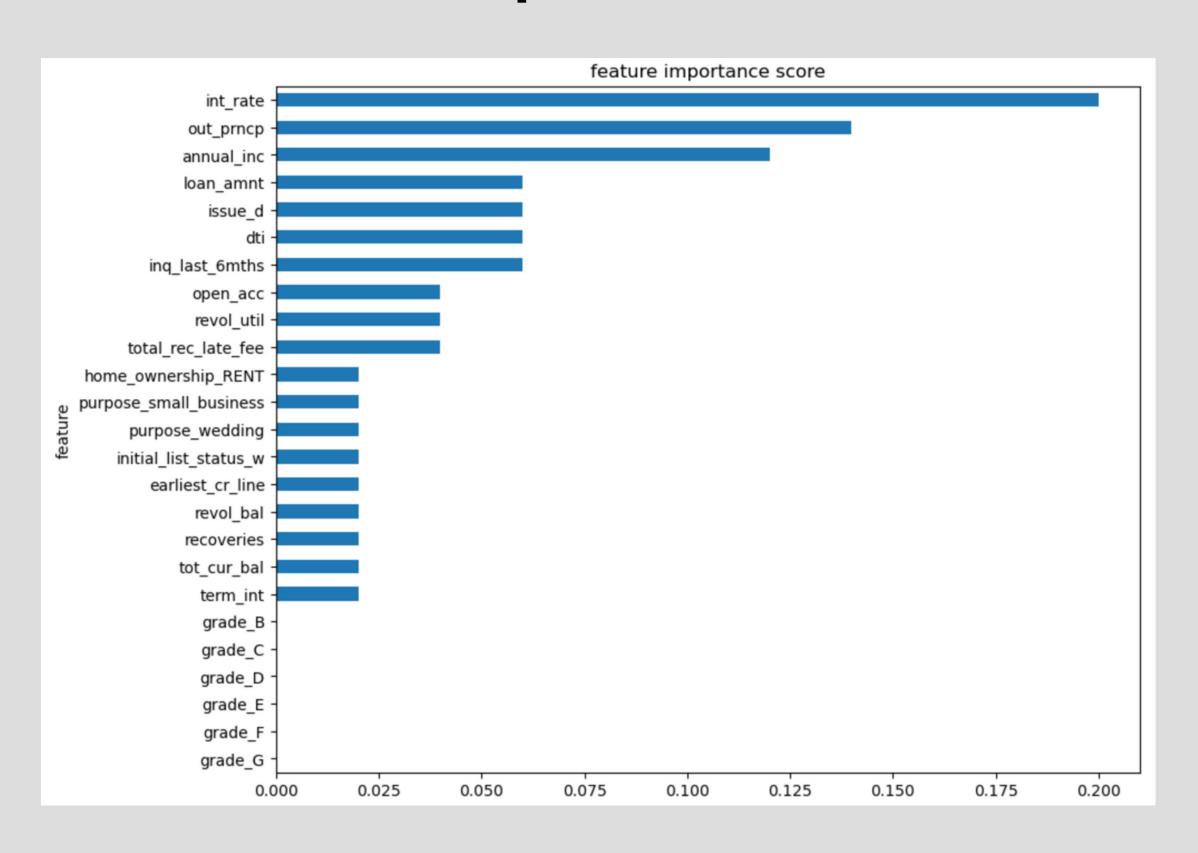
MODELLING

A lara vitla na	Evaluation Model				
Algorithm	Accuracy	Precision	Recall	AUC	AUC (Crossval)
GradientBoost	0.94	0.99	0.49	0.87	0.75
AdaBoost	0.94	0.98	0.49	0.87	0.75
XGBoost	0.94	0.97	0.52	0.91	0.70
DecisionTree	0.90	0.53	0.58	0.76	0.61

- Metrics evaluasi yang digunakan adalah AUC.
- AdaBoost memiliki algoritma yang paling baik dengan gap antara AUC train dan test yang sangat kecil dengan hasil cross validation yang paling tinggi dari yang lainnya.

FEATURE IMPORTANCE

TOP Feature Importance From AdaBoost





BUSINESS RECOMMENDATION

Feature Importance

Focusing on feature importance could be a valuable business opportunity. By identifying the most important features and building a focused and effective model, this could help to reduce credit risk and make more informed lending decisions.

Outstanding Principal

Focus monitoring more on loans that have a high remaining principal amount compared to the total amount funded. Borrowers with a high remaining principal may experience difficulty repaying the loan, and this can be a sign of a higher risk of becoming a bad loan

Interest Rate

It would be beneficial to offer flexible interest rates that can be adjusted based on the borrower's creditworthiness and payment history. This allows for a more personalized approach to lending and can help the lender manage their credit risk more effectively.

