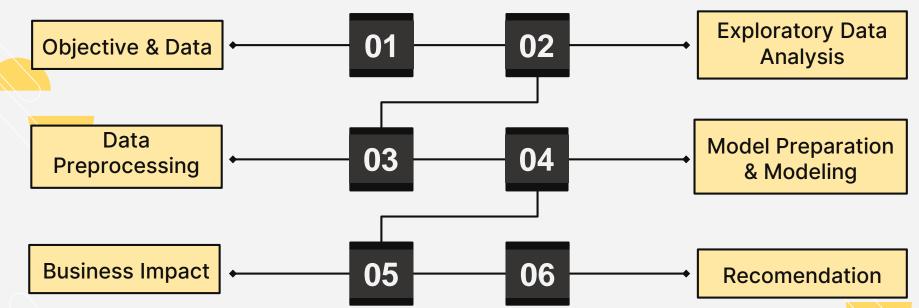


Predicting Loan Default Using Various Machine Learning Algorithms

By Dery Saefudin

Outline





Business Understanding



Loan Default is a *PROBLEM* for Financial Industry.

The failure of a debtors to repay a loan according to the terms and conditions outlined in the loan agreement.





Objective

Current

6.188 debtors are current



22%

Default

22.313 debtors are default



78%

Obtain Insight of Loan Status

Create Machine Learning Prediction

Data

Before Cleaning

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 32581 entries, 0 to 32580
Data columns (total 12 columns):
                                Non-Null Count
    Column
                                               Dtype
                                32581 non-null
                                               int64
    person_age
    person income
                                32581 non-null
                                               int64
    person_home_ownership
                                32581 non-null
                                               object
    person emp length
                                31686 non-null float64
    loan_intent
                                               object
                                32581 non-null
    loan grade
                                32581 non-null
                                               object
    loan amnt
                                32581 non-null
                                               int64
    loan int rate
                                29465 non-null float64
    loan status
                               32581 non-null int64
    loan percent income
                             32581 non-null float64
    cb person default on file 32581 non-null object
11 cb_person_cred_hist_length 32581 non-null int64
dtypes: float64(3), int64(5), object(4)
```

 Retrieved from https://www.kaggle.com/datasets/laots e/credit-risk-dataset/data

memory usage: 3.0+ MB

After Cleaning

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 28501 entries, 0 to 32580
Data columns (total 12 columns):
    Column
                                Non-Null Count
                                                Dtype
    person age
                                28501 non-null
                                                int64
    person income
                                28501 non-null
                                                int64
    person home ownership
                                28501 non-null
                                                obiect
    person emp length
                                28501 non-null float64
    loan intent
                                28501 non-null object
    loan grade
                                28501 non-null
                                                object
    loan amnt
                                28501 non-null
                                                int64
    loan int rate
                                28501 non-null float64
    loan status
                                28501 non-null int64
    loan percent income
                                28501 non-null float64
    cb_person_default_on_file
                                28501 non-null
                                                object
    cb person cred hist length
                                28501 non-null int64
dtypes: float64(3), int64(5), object(4)
memory usage: 2.8+ MB
```

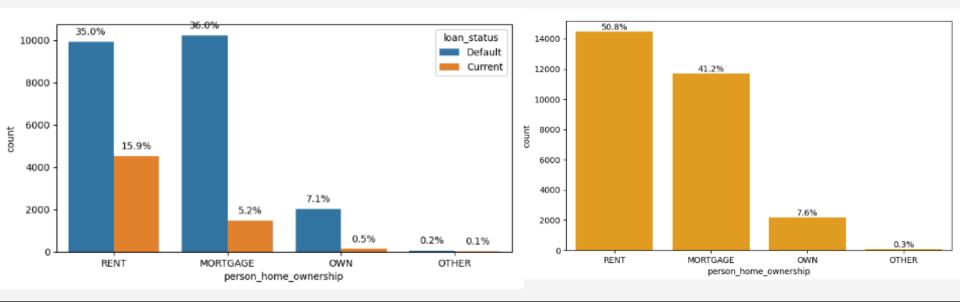
- The data has at most 32,581 rows and 12 columns before cleaning
- There are 28,501 rows and 12 columns after cleaning



Exploratory DataAnalysis

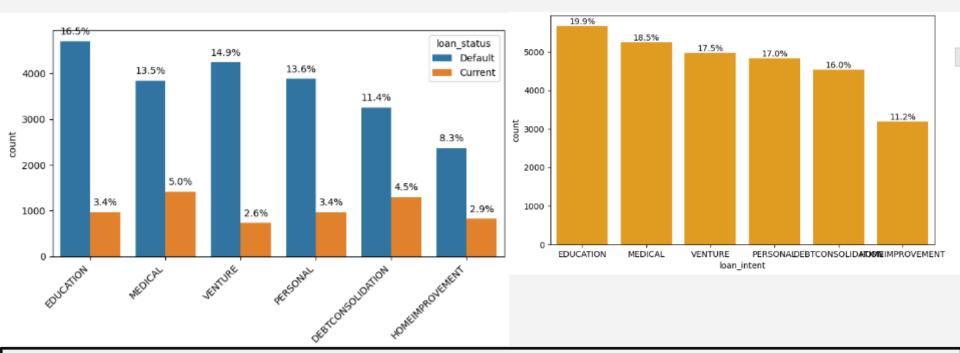


Home Ownership by Loan Status



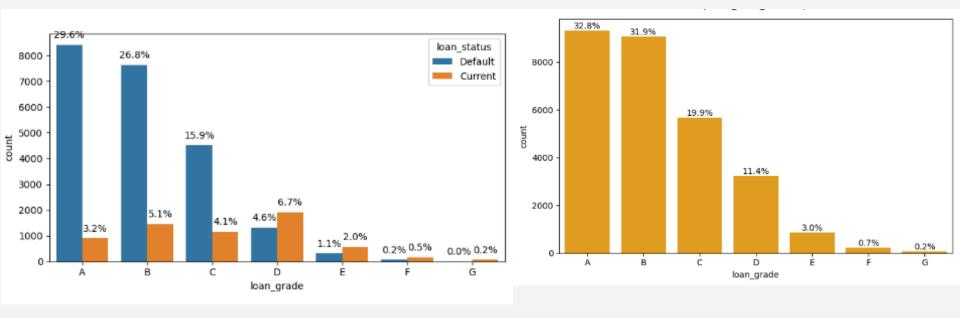
- Most of Debtors' home ownership are Rent (50.8%)
- Debtors who default the most are Occupier (15.9%)

Loan Intent by Loan Status



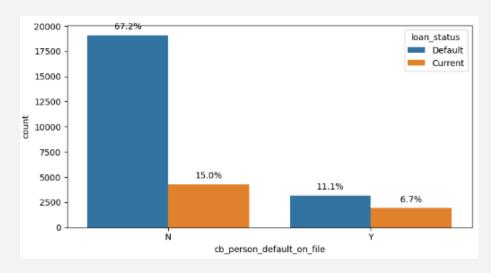
- The most common purpose of loan is Education (19.9%).
- Debtors who the most default are those with Education as the purpose (16.5%)

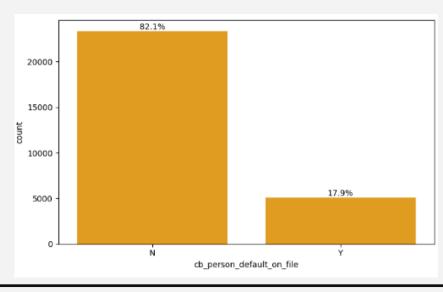
Loan Grade by Loan Status



- The most common loan grade is Grade A (32.8%).
- Debtors who the most default are those with Loan Grade A (29.6%)

Default History by Loan Status

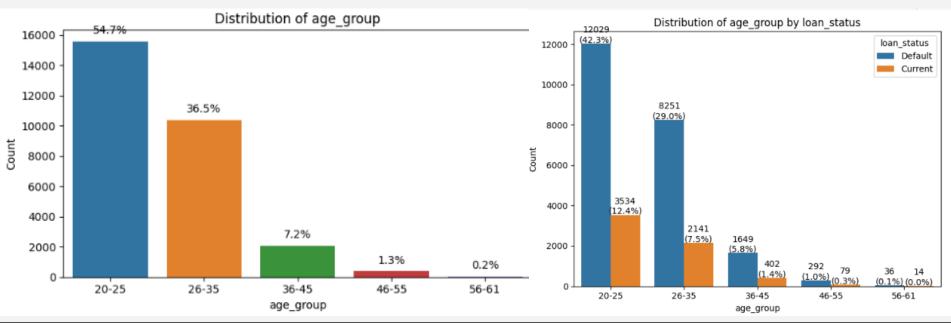




- The most debtors do not have loan default history (82.1%)
- Debtors who the most default are those without loan default history (67.2%)



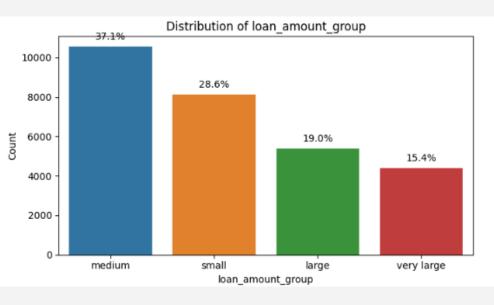
Age Group by Loan Status

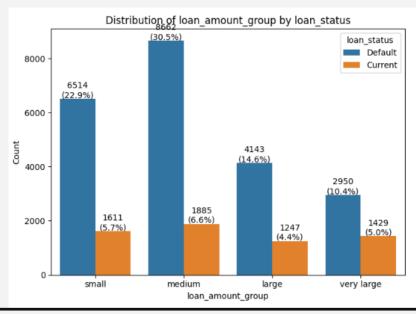


- The mostdebtors fall within the 20-25 age range (54.7%)
- Debtors who the most default are those within the 20-25 age range (42.3%)



Loan Amount Group by Loan Status

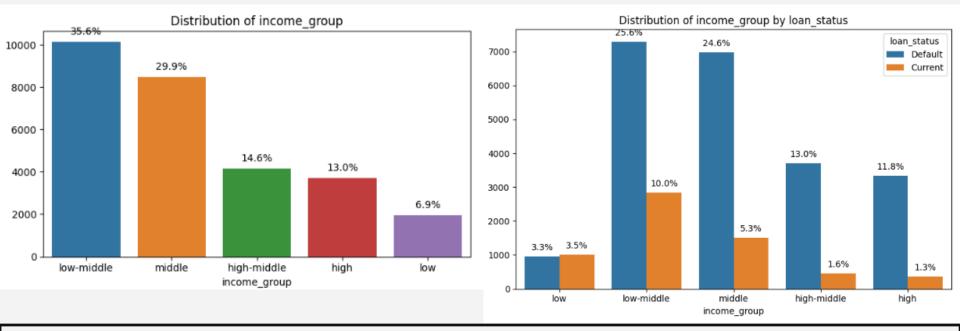




- The most common debtors are those with loan in the range of 5,000 10,000 / medium (37.1%).
- Debtors who the most default are those with loan in the range of 5,000 - 10,000 / medium (22.9%)



Income Group by Loan Status



- The most common debtors are those with income in the range of 5,000 10,000 / medium (35.6%).
 - Debtors who the most default are those with loan in the range of 5,000 10,000 / medium (25.6%)

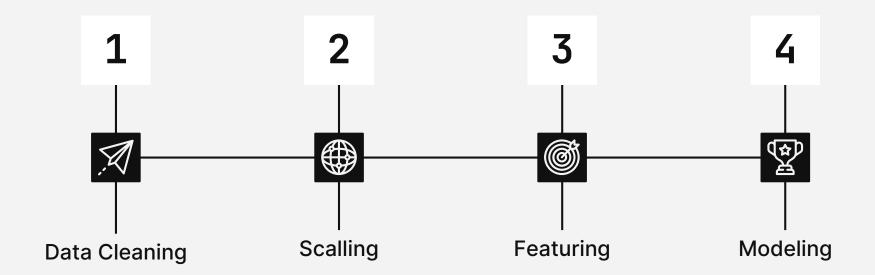
Features Engineering

After Encoding

| RangeIndex: 28441 entries, 0 to 28440 Data columns (total 44 columns): # Column Non-Null Count Dtype Operson_age 28441 non-null int64 1 person_income 28441 non-null int64 2 person_emp_length 28441 non-null int64 4 loan_int_rate 28441 non-null int64 5 loan_status 28441 non-null int64 6 loan_percent_income 28441 non-null int64 8 loan_to_income_ratio 28441 non-null float64 8 loan_to_income_ratio 28441 non-null float64 10 int_rate to_loan_amt_ratio 28441 non-null int64 10 int_rate_to_loan_emratio 28441 non-null float64 10 int_rate_to_loan_emratio 28441 non-null int64 10 int_rate_to_loan_emratio 28441 non-null int64 10 int_rate_to_loan_emratio 28441 non-null int64 11 person_home_ownership_MORTGAGE 28441 non-null int64 12 person_home_ownership_MORTGAGE 28441 non-null int64 13 person_home_ownership_MORTGAGE 28441 non-null int64 14 person_home_ownership_MORTGAGE 28441 non-null int64 15 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_EDUCATION 28441 non-null int64 17 columns (total 44 columns): 22 loan_grade_B 23 loan_grade_C 28441 non-null int64 26 loan_grade_E 28441 non-null int64 27 loan_grade_G 28 cb_person_default_on_file_N 28441 non-null int64 29 cb_person_default_on_file_Y 28441 non-null int64 29 cloan_grade_E 28441 non-null i | | | | | | | | | |
|--|--|-------------------|-------------------|---------|-------|-----------------------------|-----------|--------|--------|
| Data columns (total 44 columns): # Column Non-Null Count Dtype 24 loan_grade_D 28441 non-null int64 person_age 28441 non-null int64 person_income 28441 non-null int64 person_emp_length 28441 non-null int64 loan_int_rate 28441 non-null int64 loan_status loan_status loan_erred_hist_length 28441 non-null int64 loan_to_income_ratio 28441 non-null float64 loan_to_emp_length loan_to_emp_length 28441 non-null float64 loan_to_emp_length_ratio 28441 non-null int64 loan_prace_d 28441 non-null int64 loan_to_emp_length_ratio 28441 non-null int64 loan_to_emp_length_ratio 28441 non-null int64 loan_emp_length_ratio 28441 non-null int64 loan_ownership_ONTGAGE 28441 non-null int64 loan_intent_boundership_Rent 28441 non | <class 'pandas.core.frame.dataframe'=""></class> | | | | | loan_grade_A | 28441 no | n-null | int64 |
| # Column Non-Null Count Dtype | RangeIndex: 28441 entries, 0 to 28440 | | | | 22 | loan_grade_B | 28441 no | n-null | int64 |
| Derson_age 28441 non-null int64 26 loan_grade_F 28441 non-null int64 27 loan_grade_G 28441 non-null int64 28441 non-null int64 29 person_emp_length 28441 non-null int64 29 person_emp_length 28441 non-null int64 29 cb_person_default_on_file_N 28441 non-null int64 29 cb_person_de | Data columns (total 44 columns): | | | | 23 | loan_grade_C | 28441 no | n-null | int64 |
| 0person_age28441 non-null int6426loan_grade_F28441 non-null int641person_income28441 non-null int6427loan_grade_G28441 non-null int642person_emp_length28441 non-null float6428cb_person_default_on_file_N28441 non-null int643loan_amnt28441 non-null float6429cb_person_default_on_file_Y28441 non-null int644loan_status28441 non-null float6430income_group_low28441 non-null int645loan_status28441 non-null int6431income_group_low-middle28441 non-null int646loan_percent_income28441 non-null int6432income_group_low-middle28441 non-null int647cb_person_cred_hist_length28441 non-null int6432income_group_middle28441 non-null int648loan_to_emp_length_ratio28441 non-null float6434income_group_high-middle28441 non-null int649loan_to_emp_length_ratio28441 non-null float6435age_group_20-2528441 non-null int6410int_rate_to_loan_amt_ratio28441 non-null int6436age_group_26-3528441 non-null int6411person_home_ownership_Other28441 non-null int6437age_group_36-4528441 non-null int6413person_home_ownership_Other28441 non-null int6438age_group_36-6528441 non-null int6414person_home_ownership_Other28441 non-null int6440loan_amount_group_medium28441 non-null int64 <t< td=""><td>#</td><td>Column</td><td>Non-Null Count</td><td>Dtype</td><td>24</td><td>loan_grade_D</td><td>28441 no</td><td>n-null</td><td>int64</td></t<> | # | Column | Non-Null Count | Dtype | 24 | loan_grade_D | 28441 no | n-null | int64 |
| person_income 28441 non-null int64 person_emp_length 28441 non-null float64 loan_amnt 28441 non-null int64 loan_int_rate 28441 non-null float64 loan_percent_income 28441 non-null int64 loan_percent_income 28441 non-null int64 loan_percent_income 28441 non-null int64 loan_percent_income 28441 non-null int64 loan_to_emp_length 28441 non-null int64 loan_to_emp_length_ratio 28441 non-null float64 loan_to_emp_length_ratio 28441 non-null float64 loin_trate_to_loan_amt_ratio 28441 non-null int64 person_home_ownership_MORTGAGE 28441 non-null int64 person_home_ownership_MORTGAGE 28441 non-null int64 person_home_ownership_OWN 28441 non-null int64 person_home_ownership_OWN 28441 non-null int64 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 loan_intent_PERSONAL 28441 non-null int64 loan_intent_P | | | | | 25 | loan_grade_E | 28441 no | n-null | int64 |
| 2 person_emp_length 28441 non-null float64 3 loan_amnt 28441 non-null int64 4 loan_int_rate 28441 non-null int64 5 loan_status 28441 non-null int64 5 loan_percent_income 28441 non-null int64 6 loan_percent_income 28441 non-null int64 7 cb_person_cred_hist_length 28441 non-null int64 8 loan_to_income_ratio 28441 non-null float64 9 loan_to_emp_length_ratio 28441 non-null float64 10 int_rate_to_loan_amt_ratio 28441 non-null int64 10 person_home_ownership_OTHER 28441 non-null int64 11 person_home_ownership_OTHER 28441 non-null int64 12 person_home_ownership_OTHER 28441 non-null int64 13 person_home_ownership_OWN 28441 non-null int64 14 person_home_ownership_RENT 28441 non-null int64 15 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 17 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 19 loan_amount_group_very large 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 19 loan_amount_group_very large 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 19 loan_inten | 0 | person_age | 28441 non-null | int64 | 26 | loan_grade_F | 28441 no | n-null | int64 |
| loan_amnt 28441 non-null int64 loan_int_rate 28441 non-null float64 loan_status 28441 non-null int64 loan_percent_income 28441 non-null int64 cb_person_cred_hist_length 28441 non-null float64 loan_to_income_ratio 28441 non-null int64 loan_to_income_group_low 28441 non-null int64 loan_to_income_group_low-middle 28441 non-null int64 loan_to_income_group_low-middle 28441 non-null int64 loan_to_income_group_low-middle 28441 non-null int64 loan_to_income_group_high-middle 28441 non-null int64 loan_group_low-middle 28441 non-null int64 loan_to_income_group_low-middle 28441 non-null int64 loan_to_income_group_high-middle 28441 non-null int64 loan_group_low-middle 28441 non-null int64 loan_to_income_group_high-middle 28441 non-null int64 loan_group_low-middle 28441 non-null int64 loan_to_income_group_high-middle 28441 non-null int64 loan_group_low-middle 28441 non-null int64 lo | 1 | person_income | 28441 non-null | int64 | 27 | loan_grade_G | 28441 no | n-null | int64 |
| 4 loan_int_rate | 2 | person_emp_length | 28441 non-null | float64 | 28 | cb person default on file N | 28441 no | n-null | int64 |
| 5 loan_status 28441 non-null int64 6 loan_percent_income 28441 non-null float64 7 cb_person_cred_hist_length 28441 non-null int64 8 loan_to_income_ratio 28441 non-null float64 9 loan_to_emp_length_ratio 28441 non-null float64 10 int_rate_to_loan_amt_ratio 28441 non-null int64 11 person_home_ownership_MORTGAGE 28441 non-null int64 12 person_home_ownership_OTHER 28441 non-null int64 13 person_home_ownership_OTHER 28441 non-null int64 14 person_home_ownership_RENT 28441 non-null int64 15 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_EDUCATION 28441 non-null int64 17 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 10 int_rate_to_loan_amt_ratio 28441 non-null int64 28441 non-null int64 32 income_group_low-middle 28441 non-null int64 33 income_group_high-middle 28441 non-null int64 34 income_group_high-middle 28441 non-null int64 35 age_group_20-25 28441 non-null int64 36 age_group_26-35 28441 non-null int64 37 age_group_36-45 28441 non-null int64 38 age_group_36-45 28441 non-null int64 39 age_group_36-45 28441 non-null int64 40 loan_amount_group_small 28441 non-null int64 41 loan_amount_group_medium 28441 non-null int64 42 loan_amount_group_medium 28441 non-null int64 43 loan_amount_group_large 28441 non-null int64 44 loan_amount_group_very large 28441 non-null int64 | 3 | _ | | | 29 | cb_person_default_on_file_Y | 28441 no | n-null | int64 |
| loan_percent_income 28441 non-null float64 cb_person_cred_hist_length 28441 non-null int64 loan_to_income_ratio 28441 non-null float64 loan_to_emp_length_ratio 28441 non-null float64 loan_to_emp_length_ratio 28441 non-null float64 loan_trate_to_loan_amt_ratio 28441 non-null float64 loan_person_home_ownership_MORTGAGE 28441 non-null int64 loan_person_home_ownership_MORTGAGE 28441 non-null int64 loan_person_home_ownership_MORTGAGE 28441 non-null int64 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 loan_intent_EDUCATION 28441 non-null int64 loan_intent_MORTGAGE 28441 non-null int64 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 loan_intent_MEDICAL 28441 non-null int64 loan_intent_PERSONAL 28441 non-null int64 loan_amount_group_wery large 28441 non-null int64 loan_amount_group_very large 28441 non-null int64 loan_amount_group_very large 28441 non-null int64 loan_intent_PERSONAL 28441 non-null int64 loan_amount_group_very large 28441 non-null int64 loan_amount_group_very large 28441 non-null int64 loan_amount_group_very large 28441 non-null int64 loan_intent_PERSONAL 28441 non-null int64 loan_intent_PERSONAL 28441 non-null int64 | | | | | 30 | income group low | 28441 no | n-null | int64 |
| 7 cb_person_cred_hist_length 28441 non-null int64 8 loan_to_income_ratio 28441 non-null float64 9 loan_to_emp_length_ratio 28441 non-null float64 10 int_rate_to_loan_amt_ratio 28441 non-null float64 11 person_home_ownership_MORTGAGE 28441 non-null int64 12 person_home_ownership_OTHER 28441 non-null int64 13 person_home_ownership_OTHER 28441 non-null int64 14 person_home_ownership_RENT 28441 non-null int64 15 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_EDUCATION 28441 non-null int64 17 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 28441 non-null int64 19 loan_amount_group_medium 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 19 loan_intent_PERSON | _ | - | | | 31 | income_group_low-middle | 28441 no | n-null | int64 |
| 8 loan_to_income_ratio 28441 non-null float64 9 loan_to_emp_length_ratio 28441 non-null float64 10 int_rate_to_loan_amt_ratio 28441 non-null int64 11 person_home_ownership_MORTGAGE 28441 non-null int64 12 person_home_ownership_OTHER 28441 non-null int64 13 person_home_ownership_OWN 28441 non-null int64 14 person_home_ownership_RENT 28441 non-null int64 15 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_EDUCATION 28441 non-null int64 17 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 28441 non-null int64 34 income_group_high 28441 non-null int64 35 age_group_26-25 28441 non-null int64 36 age_group_36-45 28441 non-null int64 37 age_group_36-45 28441 non-null int64 38 age_group_46-55 28441 non-null int64 40 loan_amount_group_small 28441 non-null int64 41 loan_amount_group_medium 28441 non-null int64 42 loan_amount_group_medium 28441 non-null int64 43 loan_amount_group_large 28441 non-null int64 44 loan_amount_group_very large 28441 non-null int64 45 loan_amount_group_very large 28441 non-null int64 | 6 | | | | 32 | income group middle | 28441 no | n-null | int64 |
| 8 loan_to_income_ratio 28441 non-null float64 9 loan_to_emp_length_ratio 28441 non-null float64 10 int_rate_to_loan_amt_ratio 28441 non-null int64 11 person_home_ownership_MORTGAGE 28441 non-null int64 12 person_home_ownership_OTHER 28441 non-null int64 13 person_home_ownership_OTHER 28441 non-null int64 14 person_home_ownership_RENT 28441 non-null int64 15 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_EDUCATION 28441 non-null int64 17 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 28441 non-null int64 34 income_group_high 28441 non-null int64 35 age_group_26-35 28441 non-null int64 36 age_group_36-45 28441 non-null int64 38 age_group_36-45 38 age_group_36-45 28441 non-null int64 49 loan_amount_group_small 28441 non-null int64 40 loan_amount_group_medium 28441 non-null int64 41 loan_amount_group_medium 28441 non-null int64 42 loan_amount_group_large 28441 non-null int64 43 loan_amount_group_very large 28441 non-null int64 44 loan_amount_group_very large 28441 non-null int64 | 7 | | | | 33 | income group high-middle | 28441 no | n-null | int64 |
| 10 | | | | | 34 | _0 !_ 0 | 28441 no | n-null | int64 |
| 11 person_home_ownership_MORTGAGE 28441 non-null int64 12 person_home_ownership_OTHER 28441 non-null int64 13 person_home_ownership_OTHER 28441 non-null int64 14 person_home_ownership_RENT 28441 non-null int64 15 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_EDUCATION 28441 non-null int64 17 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 10 loan_intent_PERSONAL 28441 non-null int64 11 loan_amount_group_medium 28441 non-null int64 12 loan_amount_group_large 28441 non-null int64 13 age_group_26-35 28441 non-null int64 38 age_group_36-45 39 age_group_36-61 28441 non-null int64 40 loan_amount_group_small 28441 non-null int64 41 loan_amount_group_medium 28441 non-null int64 42 loan_amount_group_large 28441 non-null int64 43 loan_amount_group_very large 28441 non-null int64 44 dtypes: float64(6), int64(38) | _ | | | | 35 | | 28441 no | n-null | int64 |
| person_home_ownership_OTHER 28441 non-null int64 37 age_group_36-45 28441 non-null int64 38 age_group_36-45 28441 non-null int64 39 age_group_36-61 28441 non-null int64 39 age_group_56-61 28441 non-null int64 40 loan_amount_group_small 28441 non-null int64 41 loan_amount_group_medium 28441 non-null int64 41 loan_amount_group_medium 28441 non-null int64 42 loan_amount_group_large 28441 non-null int64 43 loan_amount_group_very large 28441 non-null int64 44 loan_amount_group_very large 28441 non-null int64 45 loan_amount_group_very large 28441 non-null int64 46 dtypes: float64(6), int64(38) | | | | | 36 | 0 _0 | 28441 no | n-null | int64 |
| 13 person_home_ownership_ONN 28441 non-null int64 14 person_home_ownership_RENT 28441 non-null int64 15 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_EDUCATION 28441 non-null int64 17 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 40 loan_amount_group_small 28441 non-null int64 41 loan_amount_group_medium 28441 non-null int64 42 loan_amount_group_large 28441 non-null int64 43 loan_amount_group_very large 28441 non-null int64 44 loan_amount_group_very large 28441 non-null int64 45 loan_amount_group_very large 28441 non-null int64 46 loan_amount_group_very large 28441 non-null int64 | | | | | 37 | 00 | 28441 no | n-null | int64 |
| 14 person_home_ownership_RENT 28441 non-null int64 15 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_EDUCATION 28441 non-null int64 17 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 10 loan_amount_group_medium 28441 non-null int64 42 loan_amount_group_large 28441 non-null int64 43 loan_amount_group_very large 28441 non-null int64 44 loan_amount_group_large 28441 non-null int64 45 loan_amount_group_very large 28441 non-null int64 46 loan_amount_group_large 28441 non-null int64 47 loan_amount_group_very large 28441 non-null int64 | | | | | 38 | | 28441 no | n-null | int64 |
| 15 loan_intent_DEBTCONSOLIDATION 28441 non-null int64 16 loan_intent_EDUCATION 28441 non-null int64 17 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 10 loan_amount_group_small 28441 non-null int64 41 loan_amount_group_medium 28441 non-null int64 42 loan_amount_group_large 28441 non-null int64 43 loan_amount_group_very large 28441 non-null int64 44 loan_amount_group_medium 28441 non-null int64 45 loan_amount_group_very large 28441 non-null int64 46 loan_amount_group_small 28441 non-null int64 47 loan_amount_group_medium 28441 non-null int64 48 loan_amount_group_small 28441 non-null int64 49 loan_amount_group_small 28441 non-null int64 40 loan_amount_group_small 28441 non-null int64 41 loan_amount_group_large 28441 non-null int64 42 loan_amount_group_intents 43 loan_amount_group_small 28441 non-null int64 44 loan_amount_group_small 28441 non-null int64 45 loan_amount_group_small 28441 non-null int64 46 loan_amount_group_small 28441 non-null int64 47 loan_amount_group_small 28441 non-null int64 48 loan_amount_group_small 28441 non-null int64 49 loan_amount_group_small 28441 non-null int64 | | | | | | 0 _0 | 28441 no | n-null | int64 |
| 16 loan_intent_EDUCATION 28441 non-null int64 17 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 10an_amount_group_medium 28441 non-null int64 42 loan_amount_group_large 28441 non-null int64 43 loan_amount_group_very large 28441 non-null int64 44 dtypes: float64(6), int64(38) | | | | | 40 | 0 _0 | 28441 no | n-null | int64 |
| 17 loan_intent_HOMEIMPROVEMENT 28441 non-null int64 42 loan_amount_group_large 28441 non-null int64 18 loan_intent_MEDICAL 28441 non-null int64 43 loan_amount_group_very large 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 dtypes: float64(6), int64(38) | | | | | | | | | int64 |
| 18 loan_intent_MEDICAL 28441 non-null int64 43 loan_amount_group_very large 28441 non-null int64 19 loan_intent_PERSONAL 28441 non-null int64 dtypes: float64(6), int64(38) | | | | | | | | | int64 |
| 19 loan_intent_PERSONAL 28441 non-null int64 dtypes: float64(6), int64(38) | | | | | | | | | |
| | | | | | | , _ , | 20.41 110 | | 1.1.00 |
| as assignment and assignment and assignment and assignment and assignment assignment as a second as a seco | | | | | | | | | |
| | | | 20.72 11011 11011 | 2 | CIIIO | . j asager 515 115 | | | |

^

Model Preparation





Modeling

Random Forest

Logistic Regression

K-Neighbour

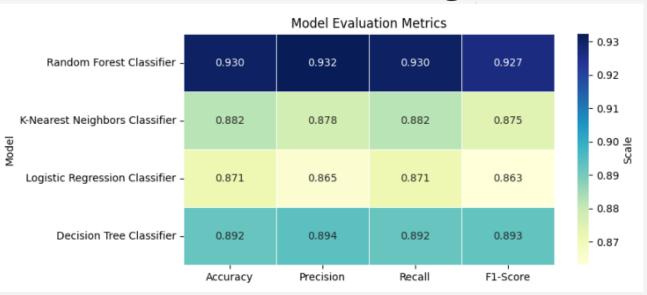
Decission Tree

Base Model

Hyperparameter Tuning (Grid Search)

Model

Before Tuning



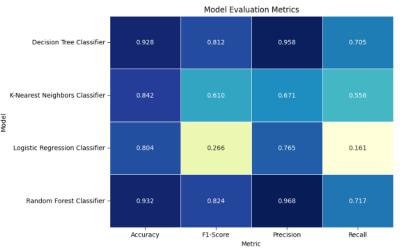


Random Forest
Precision (0.932)
Recall (0.892)





After Tuning



96 %

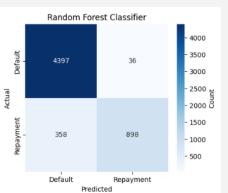


0.4

- 0.2

Model can predict the true default debtors correctly

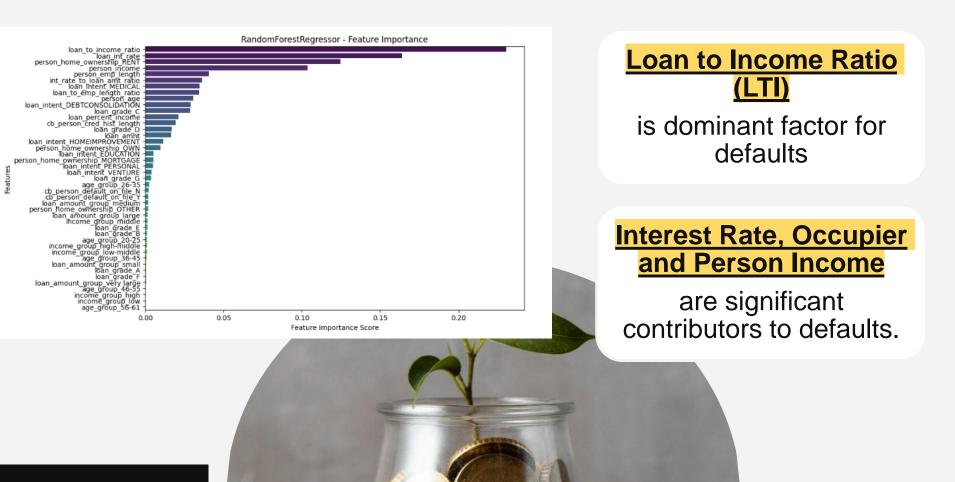




Business Impact

- The model can correctly predict 96 of the 100 default case and 29 default case can not be predicted correctly.
- If on average the bank's loss due to the default case is Rp 100 million per case & month, then using this model, the total potential losses that could be prevented are Rp 4.2 billion.

Feature Important





Business Recomendation

Improved risk management strategies:

- Adjusting loan principal limits (LTI < 35%)
- Adjusting interest rates
- Requesting additional documentation
- Adding additional approval processes
- Targeting Low Risk Market Segment





