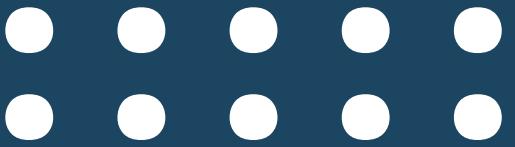


Credit Risk Loan Predictions

ADIFTA WISNU WARDANA



Problem & Goals

Companies engaged in lending money have data on all customer transactions from 2007 to 2014. In this case, not all customers/borrowers have the status of successfully returning/paying loan bills that should have been deposited with the company.

With advances in information technology in the field of data science, companies want to have a system that is able to operate in determining the predictions of their customers who have a risk of default so that this can help companies overcome losses that can occur.



Table of Contents



f

Data Exploration

Data Preparation

Data Modeling

Result





Data Exploration

The data used to create machine learning models is a dataset of loan records from 2007 to 2014. The dataset consists of 466,285 rows and 75 columns consisting of categorical and numerical data.



Data Preparation

Cleansing

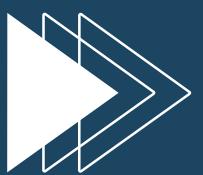
- Drop features with null values > 25%
- Define target predict labels from the data
- Impute missing value from numerical features with a median
- Processing outliers from numerical features
- Drop unnecessary feature

Feature Engineering

- Encode all categorical feature with one hot encoding
- Split data into train and test data
- Perform feature scaling on numerical data
- Process the target data imbalance that will be predicted using the smote method

Data Modeling

To build a system capable of predicting credit risk, we prefer to use logistic regression by using 80 percent of the data as material for machine learning and the remaining 20 percent as test data to test the ability of the system to predict credit risk.



Result

By using logistic regression algorithm to create a machine learning model for predicting credit risk, the following results are obtained.

Finally, it can be said that the developed method is quite accurate at predicting credit risk, which is of course very helpful for businesses to screen potential customers to reduce the danger of loss from clients who won't pay.

ACCURACY

98.62%

PRECISION

98.82%

RECALL

99.46%

AUC

99.42%



THANK YOU!

Reach me at :

adiftawisnu818@gmail.com

<https://www.linkedin.com/in/adifta-wisnu-wardana/>