

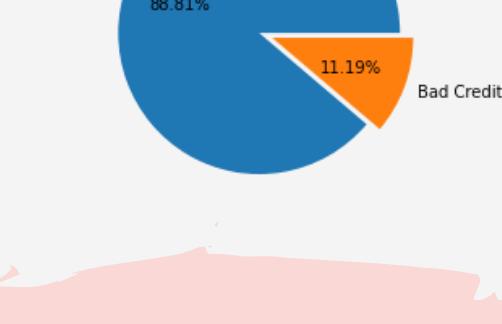
ABOUT THE LOANERS PREDICT

Infographic for Final project of ID/X Partners Data Scientist
Virtual Internship Program, Rakamin Academy
By Isaac Dwadattuyah Haikal Azziz

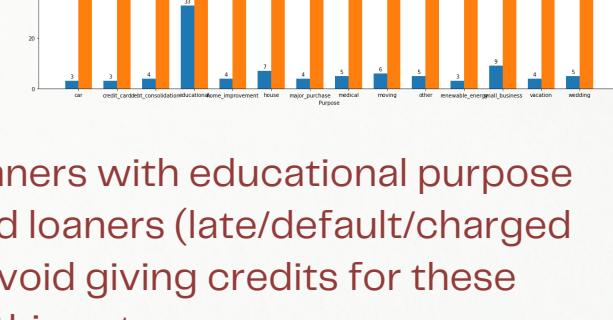
GOOD CREDIT VS BAD CREDIT

It's important to predict whether credit applicants fall into the category of good or bad when given the loan. Company lose about \$394.351.677 alone from bad credit applicants (11,19% cases)

Good Credit and Bad Credit Comparison



CAN WE PREDICT THAT?



Grade G loan applicant and loaners with educational purpose have higher chance to be a bad loaners (late/default/charged off). Its recommended to avoid giving credits for these applicant within category

MODEL TO PREDICT

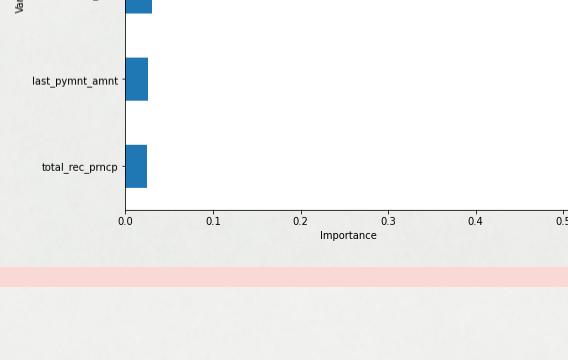
Model	Resampling Method	F1-Score Test
Decision Tree	Without Resampling	0.9890
	Undersampling	0.8649
	Oversampling	0.9923
Random Forest	Without Resampling	0.9941
	Undersampling	0.9006
	Oversampling	0.9994

I tried 2 kind of model combined with 3 resampling method, Using F1-Score as goodness metric, we can conclude the best model was random forest with oversampling method

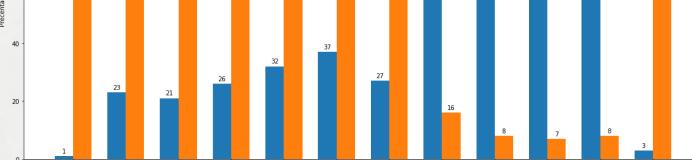
WE ALSO GET MOST IMPORTANT VARIABLE

Month of the last payment was the most important variable in prediction model

VARIABLE



WHY MONTH OF LAST PAYMENT?



Well, Looks like the last-payer in august-november has higher chance to be categorized as bad loaners. I dont know the cause tho.