APPLICATION OF BUSINESS ANALYTICS IN CONSUMER (RETAIL) CREDIT RISK

Andreas Zaras, Data Scientist

REFERENCES

- Anderson, Billie, and J.M. Hardin, 2009. *Development of Credit Scoring Applications Using SAS® Enterprise Miner Course Notes.* Cary, NC: SAS Institute Inc.

- Peter Christie et al, 2011. *Applied Analytics Using SAS Enterprise Miner Course Notes*. Cary: SAS Institute Inc.

- SAS Institute Inc. *Building Credit Scorecards Using Credit Scoring for SAS® Enterprise Miner: A SAS Best Practices Paper*. Cary, NC.

- SAS Institute Inc. *Building Consumer Credit Scoring Models with Enterprise Miner: A SAS Best Practices Paper*. Cary, NC.

THE OLD CREDIT RISK MANAGEMENT PROCESS

A few decades ago the retail credit risk management process was based on an interview, so it was judgmental and hence subjective.

Today the retail credit risk management process is based on data, analytics and technology and hence it is objective.

Consumer (or retail) credit risk is the risk of loss due to a customer's default on a consumer credit product, such as a mortgage, unsecured personal loan, credit card, overdraft etc.

Consumer (or retail) credit risk management is a set of processes dedicated to the measurement, prediction and control of losses due to retail credit risk

THE TOOL FOR RETAIL CREDIT RISK MANAGEMENT

A credit scoring model is a statistical tool that is typically used in supporting organizations when making credit decisions

Types of Credit Scoring Tools

Application Credit Scoring

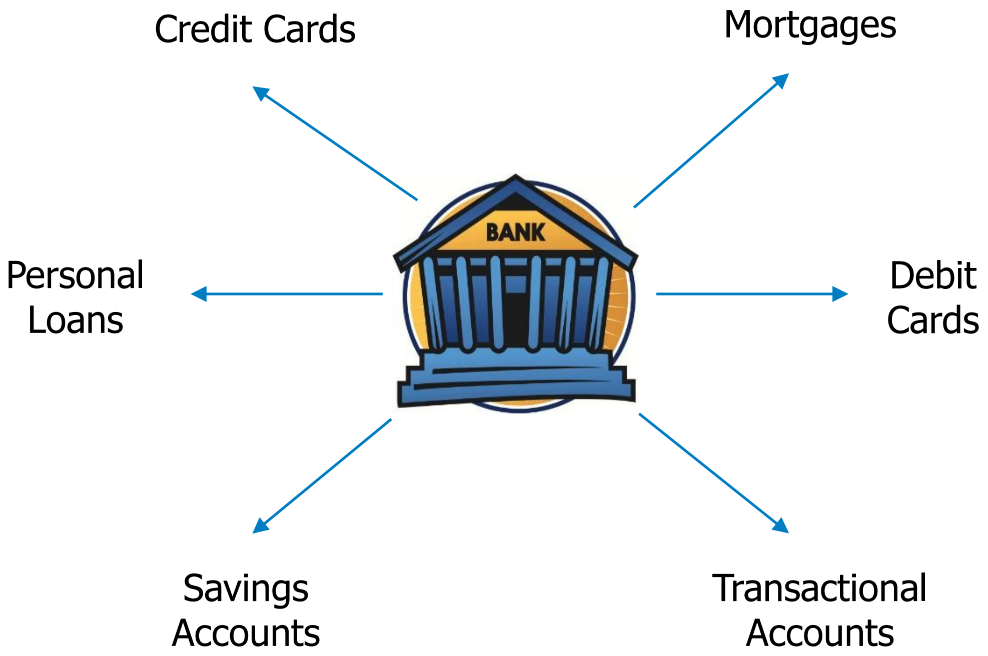
Behavioral Credit Scoring

Collections Credit Scoring

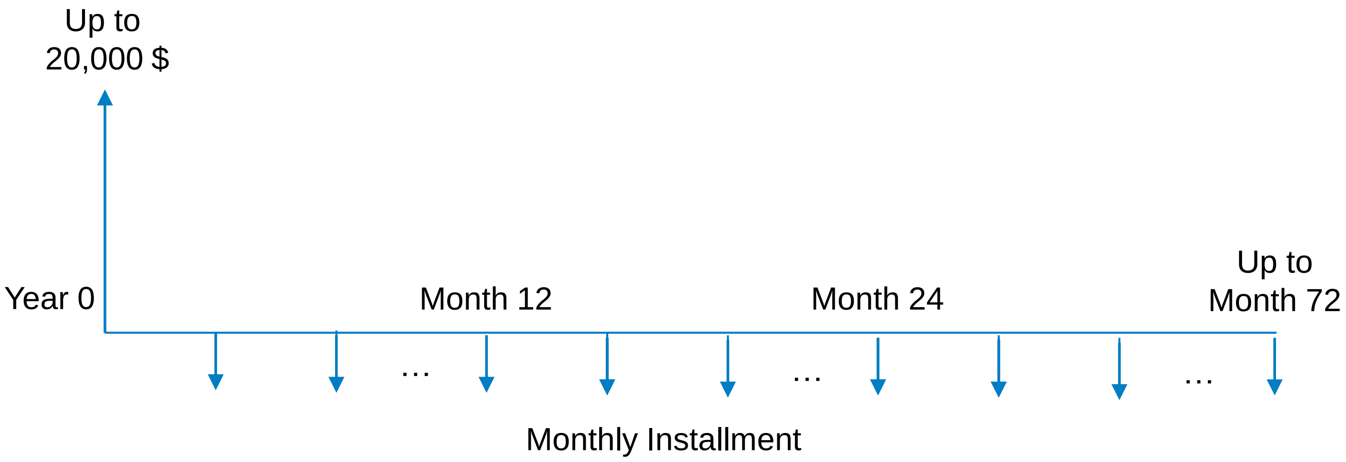
THE CASE OF COOPERATIVE POPULAR BANK (CPB)

CPB is a retail bank i.e. it offers banking services to individuals and small business.

In this case study we focus on consumer loans up to $20,000 for which the repayments can be spread from 1 to 6 years i.e. Small Consumer Loans (SCL).

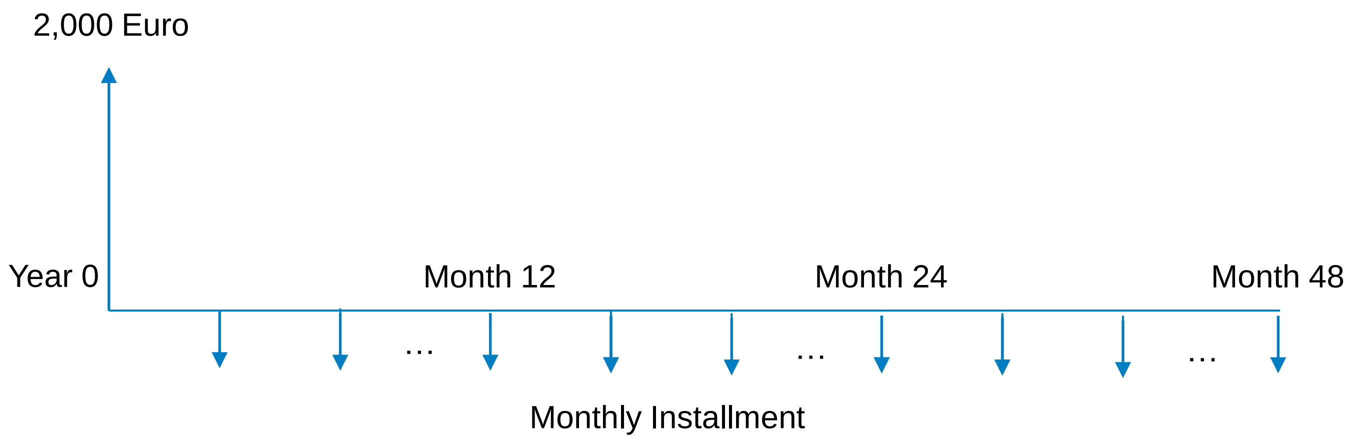


THE CASE OF COOPERATIVE POPULAR BANK (CPB)



Interest Rate = 12%

THE CASE OF COOPERATIVE POPULAR BANK (CPB)



Principal Amount of Loan = 2,000 Euro

Repayment Period = 48 months

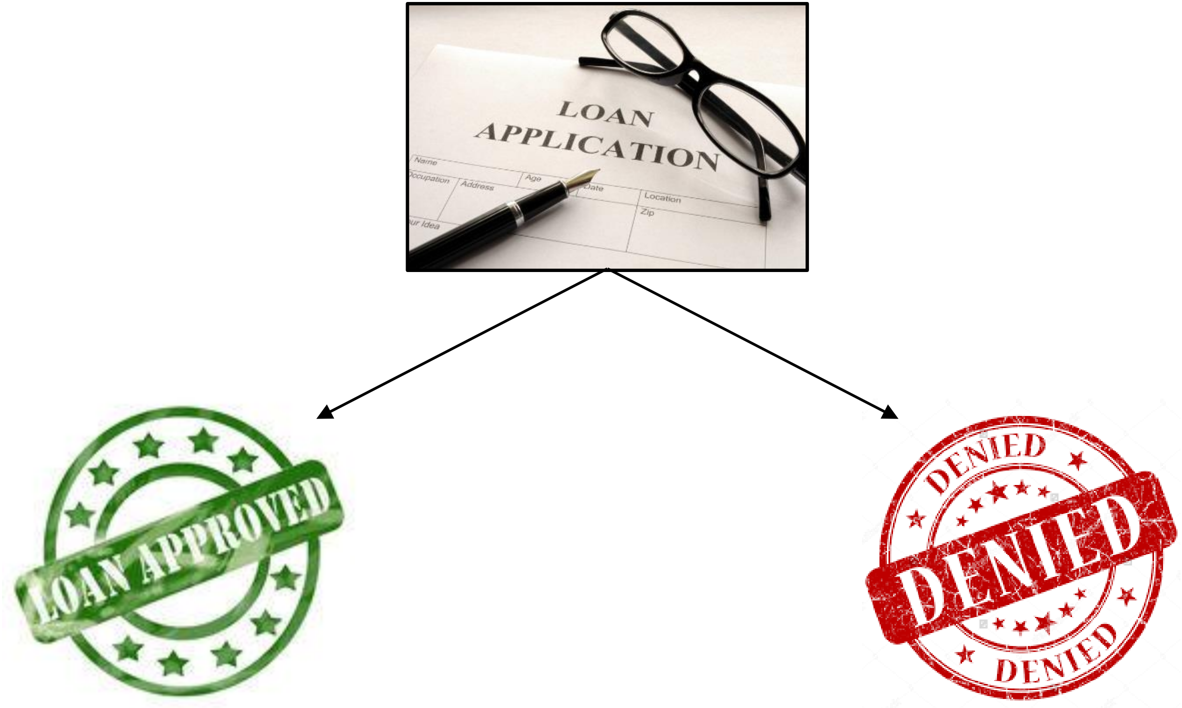
Interest Rate = 12%

Monthly Installment= 44.27 Euro

Total Repayment (Principal amount + Total Interest) = 44.27 * 48 = 2,124.9 Euro

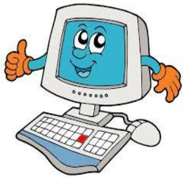
Bank’s Profit = Total Interest = 2,124.9 – 2,000 = 124.9 Euro

THE CASE OF COOPERATIVE POPULAR BANK (CPB)



DEFINITION OF APPLICATION CREDIT SCORING

An application credit scoring model is a statistical tool that is typically used in the decision-making process of accepting. or rejecting a loan.



Information about the applicant (historical data)

Probability of Being a Bad Customer

p = x%

DEFINITION OF CREDIT SCORING (CONT'D)

A credit scoring model is a statistical tool that is typically used in the decision-making process of accepting. or rejecting a loan.

Probability of Being
a Bad Customer

p = x%

If p > Cut – Off ⇨ Reject the Application

If p < Cut – Off ⇨ Accept the Application

Objective of the Credit Risk Department

Develop an Application Credit Scoring tool for the Small Consumer Loans (SCL).

Steps to Follow

Collect historical data

Analyze the data and develop the statistical model

Apply the model to new applicants

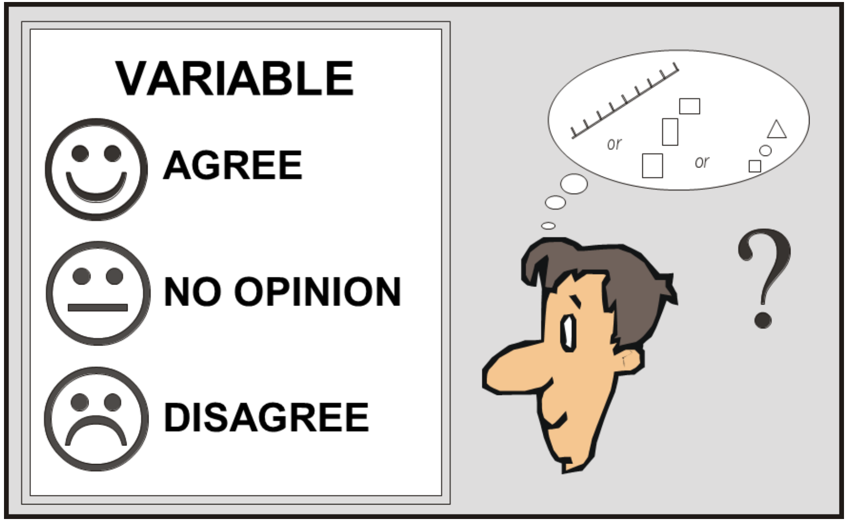
HISTORICAL DATA ABOUT 1,000 PAST BORROWERS

25 Variables

1,000 Past Borrowers

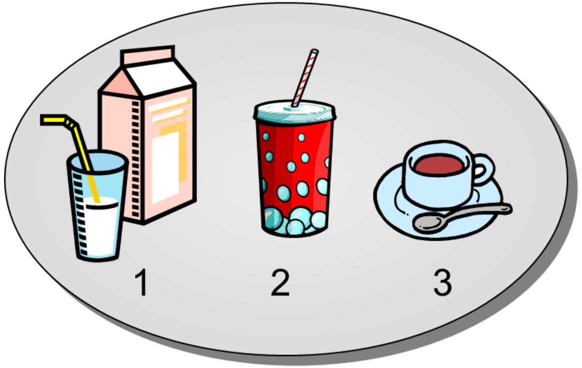
ID	Customer Attributes							Good/ Bad (1 / 0)	Principal Amount Unpaid	Total Interest	Principal Loan + Total Interest
	Age	Amount	Checking	Duration	……	Employed	Foreign				
100	35	1169	1	6	…..	5	1	0	0	83.18	1252.18
250	45	5951	2	48	…..	3	0	1	4761	371.626	6322.626
800	60	2096	2	12	…..	4	0	0	0	138.725	2234.725
…..	…..	…..	…..	…..	…..	…..	…..	…..	…..	…..	…..
1000	29	4870	1	24	…..	1	1	1	3896	310.1929	5180.193

IDENTIFYING THE SCALE OF MEASUREMENT



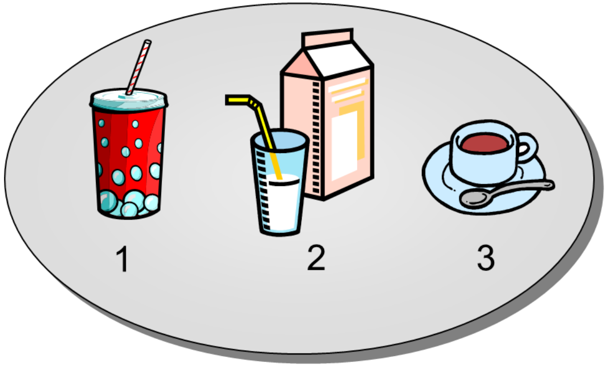
Before analyzing, identify the measurement scale for each variable (continuous, nominal, or ordinal).

NOMINAL VARIABLES



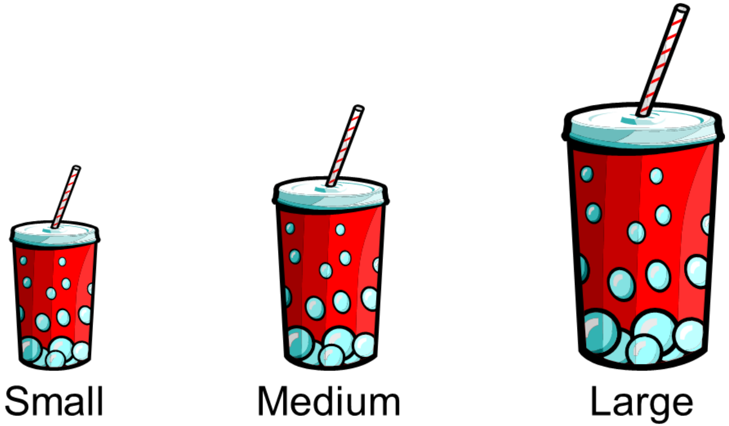
Variable: Type of Beverage

Or



Nominal variables have values with no logical ordering.

Variable: Size of Beverage



Ordinal variables have values with a logical order.

However the relative distances between the values are not clear.

AVAILABLE HISTORICAL DATA (25 VARIABLES)

Age	Age in years	Interval
Amount	Amount of loan	Interval
Checking	Status of existing checking account: 1: No Checking Account, 2: <$0, 3: 0 - <$200, 4: >=$200	Ordinal
Coapp	Other applicants/ guarantors: 1: none, 2: co-applicant, 3: guarantor	Nominal
Depends	Number of dependents	Interval
Duration	Duration in months	Interval
Employed	Presently employed since: 1: unemployed, 2: <1year, 3: 1 to <4 years, 4: 4 to <7 years, 5: >=7 years	Ordinal
Existcr	Number of existing credits at this bank	Interval

AVAILABLE HISTORICAL DATA (25 VARIABLES)

Foreign	Foreign worker: 1: yes, 2: no	Binary
Housing	1: rent, 2: own, 3: for free	Nominal
Installp	Installment rate in percentage of disposable income	Interval
Job	1: Unemployed/ Unskilled - non resident, 2: Unskilled - resident, 3: Skilled employee/ official, 4: Manager/ Shelf Employed/ Highly Qualified employee officer	Nominal
Marital	Marital Status: 1: Male: Divorced/ Separated, 2: Female: Divorced/ Separated/ Married 3: Male: Single, 4: Male married/ windowed, 5: Female: Single	Nominal
Purpose	Purpose of loan: 0: Car (new), 1: Car (used), 2: Furniture/ Equipment, 3: Radio/ TV, 4: Domestic appliances, 5: House Repairs, 6: Education, 7: Vacation, 8: Retraining	Nominal
Other Loan Obligations	1: Bank, 2: Stores, 3: None	Nominal

AVAILABLE HISTORICAL DATA (25 VARIABLES)

Savings	Savings Account: 1: <$100, 2: 100-<$500, 3: $500-<$1000, 4: >= $1000,	Ordinal
Telephone	1: None, 2: Yes, registered under the customer's name	Binary
Credit Cards in Other Banks	1: No Credit Cards, 2: 1 Credit Card, 3: 2 Credit Cards, 4: >= 2 Credit Cards	Ordinal
Good/ Bad	Good/ Bad payer	Binary
Amount Unpaid of Principal Loan	Amount unpaid of principal loan	Interval
Monthly Interest	Monthly interest	Interval
Principal Loan + Total Interest	Principal Loan + Total Interest	Interval
Bank's Profit = Total interest	Total Interest	Interval

AVAILABLE HISTORICAL DATA (25 VARIABLES)

Property	1: Real Estate, 2: Car, 3: Life Insurance, 4: No Property	Nominal
Resident	Years beginning permanent residence	Interval
History	0: No Credits Taken/ All Credits Paid Back Dully; 1: All Credits at this Bank Paid Back Dully; 2: Existing Credits Paid Back Dully Until Now; 3: Delay in Paying Off in the Past; 4: Critical Account/ Other Credits Existing (Not at this Bank)	Nominal

AVAILABLE HISTORICAL DATA (25 VARIABLES)

ID	Customer Attributes							Good/ Bad (0 / 1)	Principal Amount Unpaid	Total Interest	Principal Loan + Total Interest
	Age	Amount	Checking	Duration	……	Employed	Foreign				
100	35	1169	1	6	…..	5	1	0	0	83.18	1252.18
250	45	5951	2	48	…..	3	0	1	4761	371.626	6322.626
800	60	2096	2	12	…..	4	0	0	0	138.725	2234.725
…..	…..	…..	…..	…..	…..	…..	…..	…..	…..	…..	…..
1000	29	4870	1	24	…..	1	1	1	3896	310.1929	5180.193

Bad/ Good	Frequency	Percent
1	301	30.10%
0	699	69.90%

INTRO TO MACHINE LARNING/ SUPERVISED LEARNING

Andreas Zaras, Data Scientist

What do Decision Trees do for us?

- They help us find which attributes or characteristics of the entities i.e. past borrowers are important for classifying good borrowers from bad borrowers.

- They help us predict which entities i.e. past borrowers will be good and bad and with what probability.

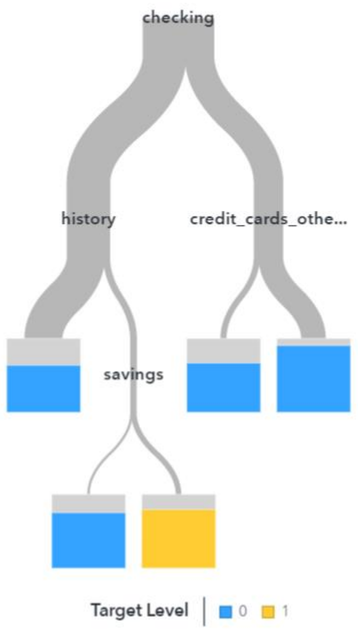
As you remember for each past borrower we have 25 characteristics e.g. age, gender, etc. The decision tree will tell us which of those characteristics are important for classifying a past borrower as a good or a bad borrower.

Available Data

(25 Variables)

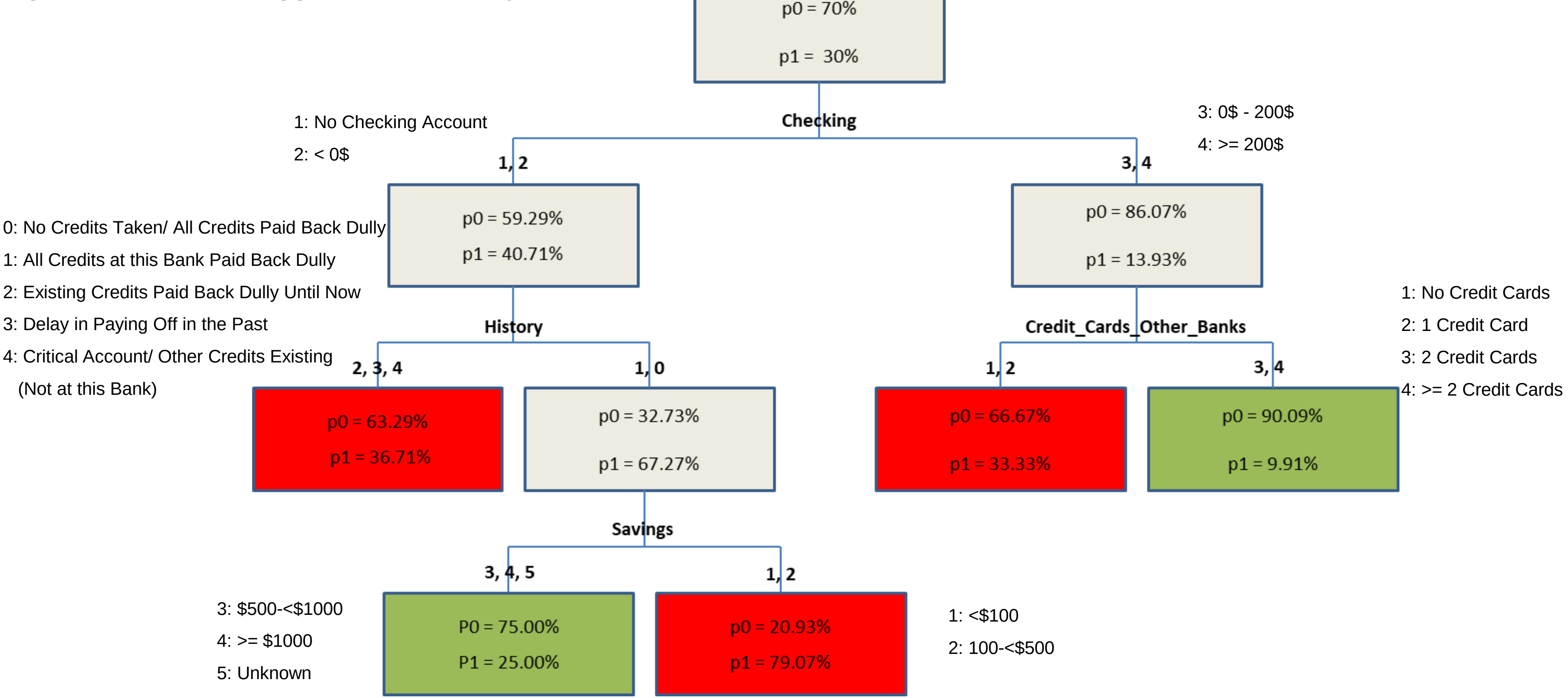
Credit Cards in Other Banks	1: No Credit Cards, 2: 1 Credit Card, 3: 2 Credit Cards, 4: >= 2 Credit Cards
Amount	Amount of loan
Checking	Status of existing checking account: 1: No Checking Account, 2: <$0, 3: 0 - <$200, 4: >=$200
Coapp	Other applicants/ guarantors: 1: none, 2: co-applicant, 3: guarantor
Property	1: Real Estate, 2: Car, 3: Life Insurance, 4: No Property
Depends	Number of dependents
Duration	Duration in months
Age	Age in years
Existcr	Number of existing credits at this bank
………..	………..
History	0: No Credits Taken/ All Credits Paid Back Dully; 1: All Credits at this Bank Paid Back Dully; 2: Existing Credits Paid Back Dully Until Now; 3: Delay in Paying Off in the Past; 4: Critical Account/ Other Credits Existing (Not at this Bank)
Other Loan Obligations	1: Bank, 2: Stores, 3: None
Purpose	Purpose of loan: 0: Car (new), 1: Car (used), 2: Furniture/ Equipment, 3: Radio/ TV, 4: Domestic appliances, 5: House Repairs, 6: Education, 7: Vacation, 8: Retraining
Foreign	Foreign worker: 1: yes, 2: no
Savings	Saving account: 1: <$100, 2: 100-<$500, 3: $500-<$1000, 4: >= $1000
Telephone	1: None, 2: Yes, registered under the customer's name
Employed	Presently employed since: 1: unemployed, 2: <1year, 3: 1 to <4 years, 4: 4 to <7 years, 5: >=7 years
InstallP	Installment rate in percentage of disposable income

The Decision Tree Model



The Decision Tree Model

If p1>= 28.46% then Applicant = Bad Payer



CREATING A PROFIT MATRIX

In the case study under consideration, there are four outcome/ action combinations:

	Grant Loan	Reject Loan
Bad Borrower		
Good Borrower		

Each of these outcome/action combinations has a profit/ negative profit (loss) consequence. Some of the profit consequences are obvious. For example, if you reject a loan to a bad borrower you do not make any profit and you don’t have any loss. For this analysis, the upper right cell can be immediately set to zero.

	Grant Loan	Reject Loan
Bad Borrower		0
Good Borrower		

...

CREATING A PROFIT MATRIX (2)

In order to reject the loan the expected profit from rejecting the loan should be greater than the expected profit from granting the loan. So:

$$0 \times p1 + (-210.2339) \times (1 - p1) > -1{,}056.848 \times p1 + (210.2339) \times (1 - p1)$$

$$-210.2339 + 210.2339 \times p1 > -1{,}056.848 \times p1 + 210.2339 - 210.2339 \times p1$$

$$-420.4678 > -1{,}477.316 \times p1$$

$$p1 > 28.46\%$$

Bad/ Good	Grant Loan	Reject Loan
Bad Borrower	-1,056.848	0
Good Borrower	210.2339	-210.2339

CONCLUSION – WRAP UP

Andreas Zaras, Data Scientist

CREDIT SCORING - CONCLUSIONS

- Over the past 30 years the credit retail industry is characterized by growing demand, fierce competition and advances in information technology.
- These characteristics have changed the way retail credit decisions are made.
- So we have passed from interview based methods and hence subjective, to data analytics based methods to decide whether to grant credit or not.
- This section demonstrated how we can use a state of the art software tool – SAS Enterprise Miner – to evaluate the creditworthiness of new customers requesting loans using analytics - based methods.

THANK YOU!