Credit One.

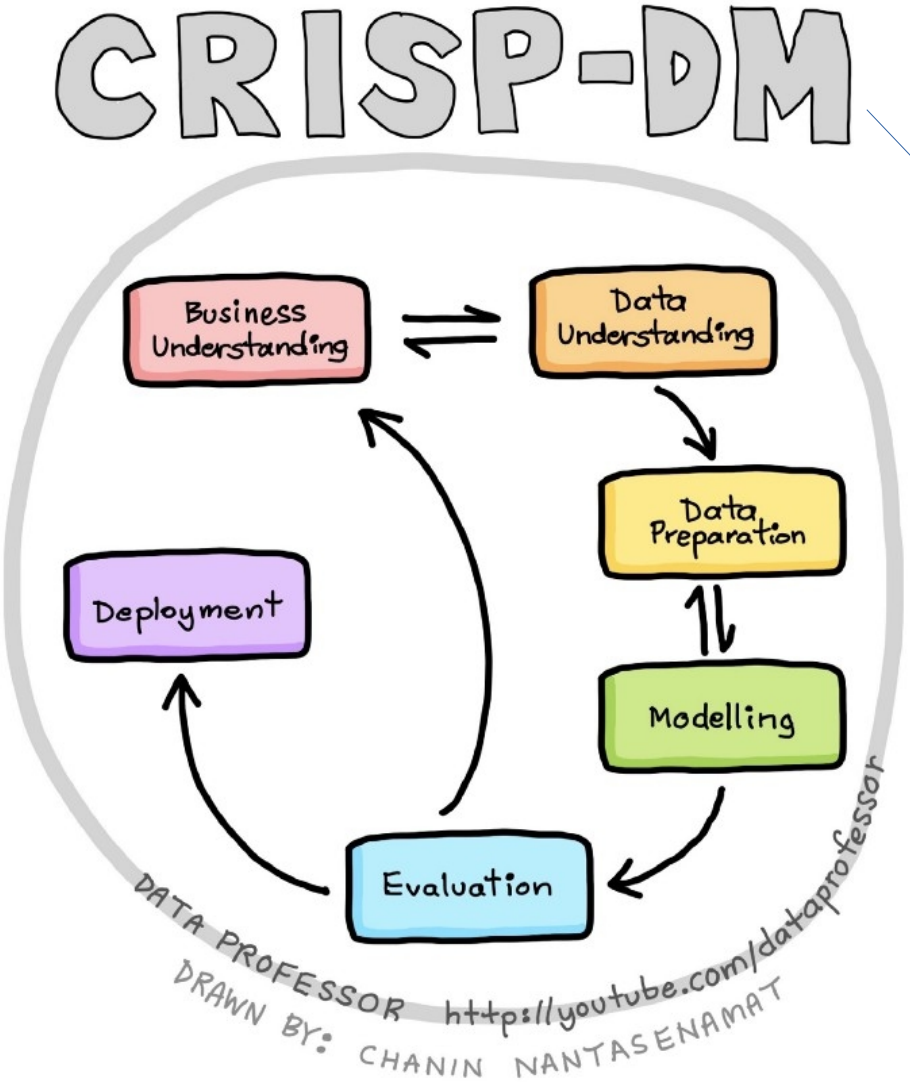
Credit Scoring Service.

Predictive Analysis.

Index

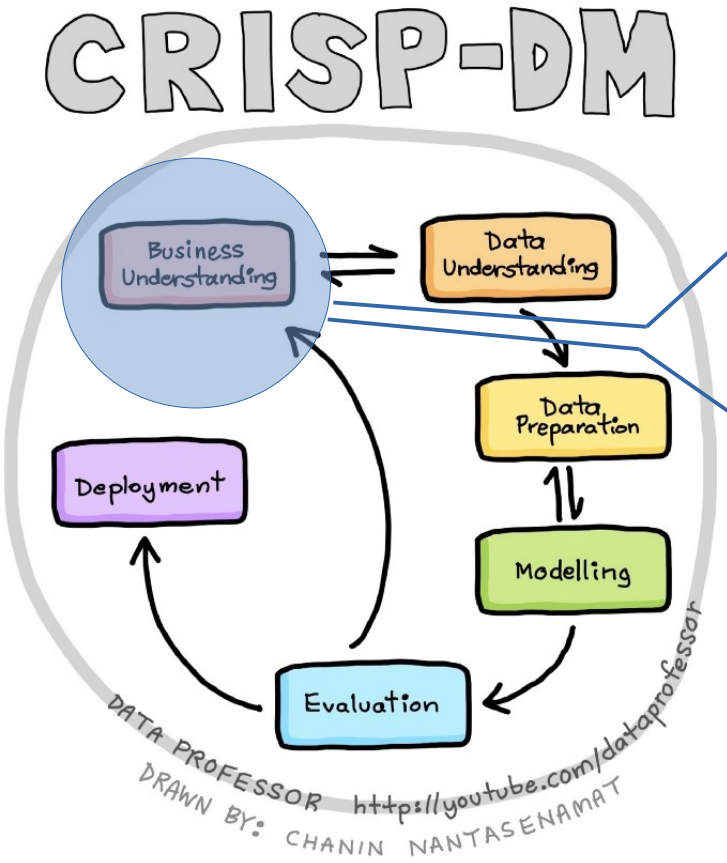
- Framework
- Some Ideas after fast overview
- Possible pitfalls

Project Framework



CRoss **I**ndustry **S**tandard **P**rocess for **D**ata **M**ining

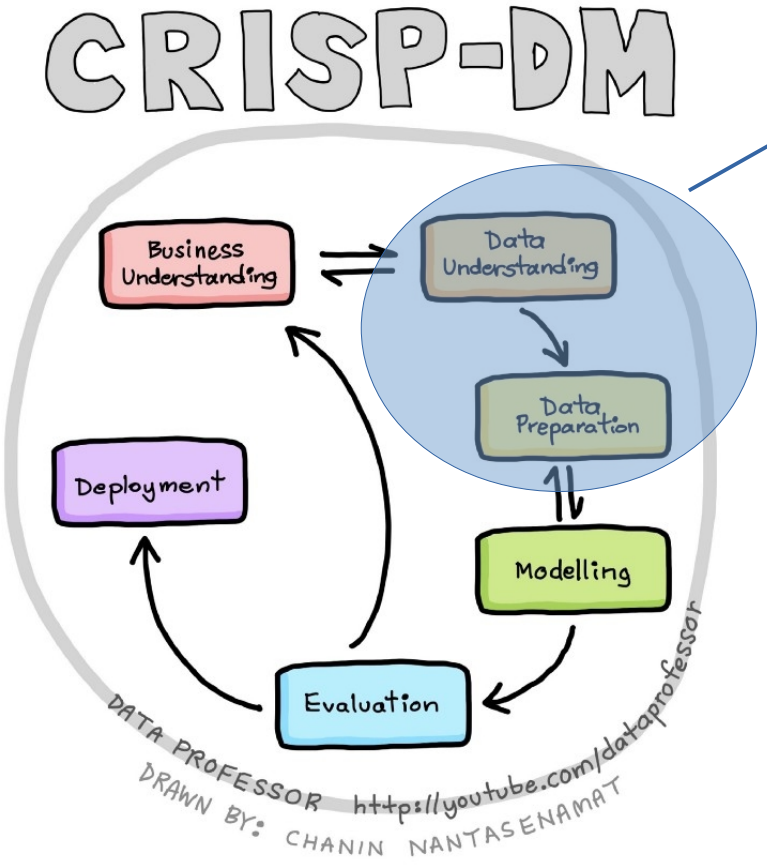
Goals



<u>Keep defaults as low as possible.</u>

Understand how much credit to allow someone to use or, at a very least, if someone should be approved or not.

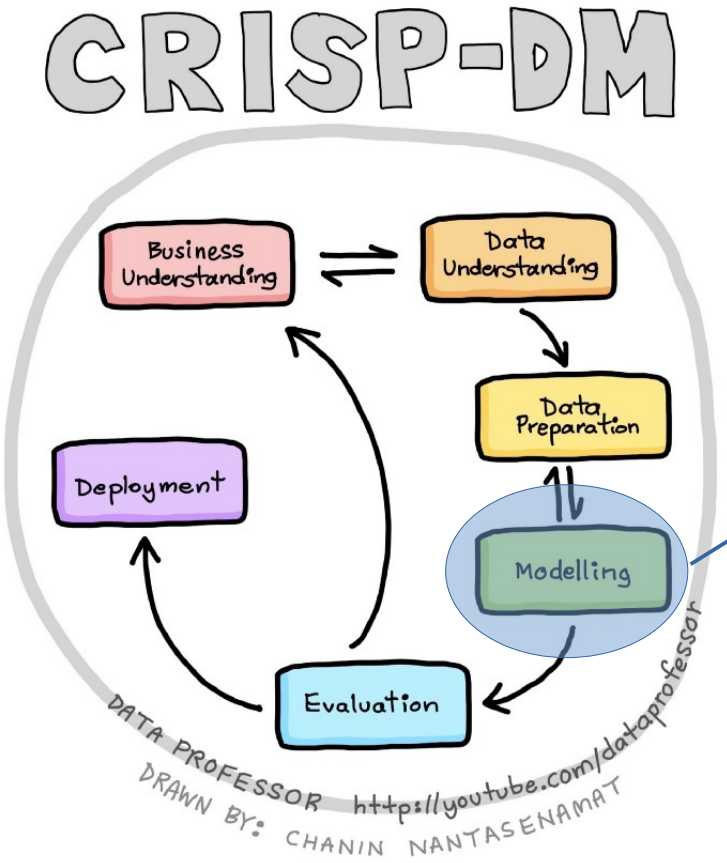
Data Understanding & preparation



Get Data
Exploratory Data Analysis
Cleaning, Relabeling, grouping, etc
Feature engineering

As required, all the project will be stored in a GitHub Account:
https://github.com/alefbenages/CreditOne

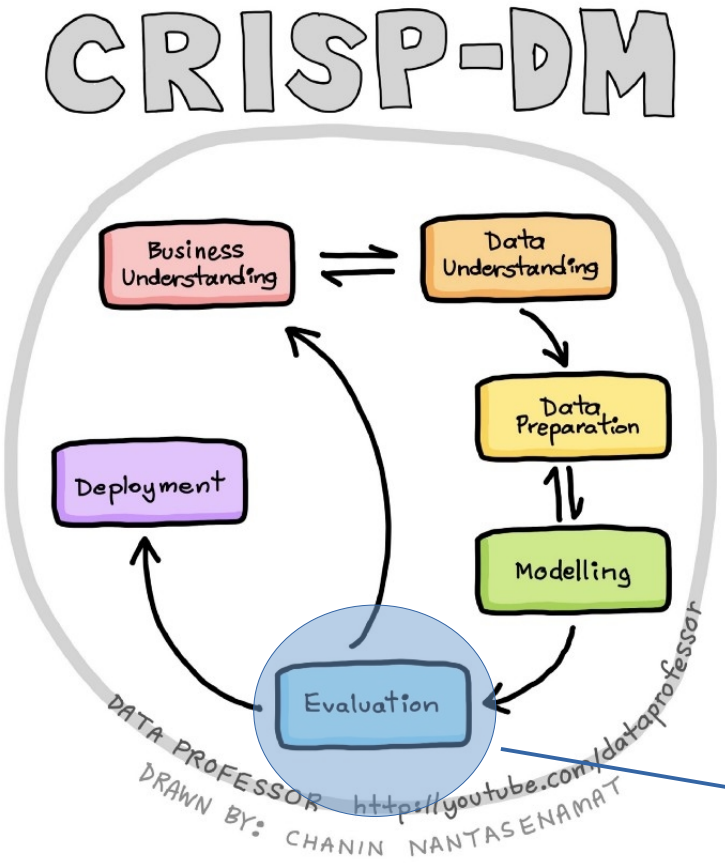
Modeling



Predictive Analysis (What will happen?)

It's a **Classification problem**.

Evaluation



Evaluate
Rinse and repeat
Future
Repeat

Some Ideas after a fast Overview

data divided in:

±30.000 rows and,

23 categories= ID, Sex, Marital Status, Education, Age, History of past payments, Amount of bill statement, Amount of previous payments, Client's behavior.

Default Probability = 22%

Some Ideas after a fast Overview

SEX → 60% Female // 40% Male

EDUCATION →

AGE →

MARITAL STATUS→

SOME GRAPHS???

Potential Pitfalls

> Some undocumented categories
  Education 0,4,5,6=Other

> Why there is no Pay_1 ??

> Some cases of mixed data types

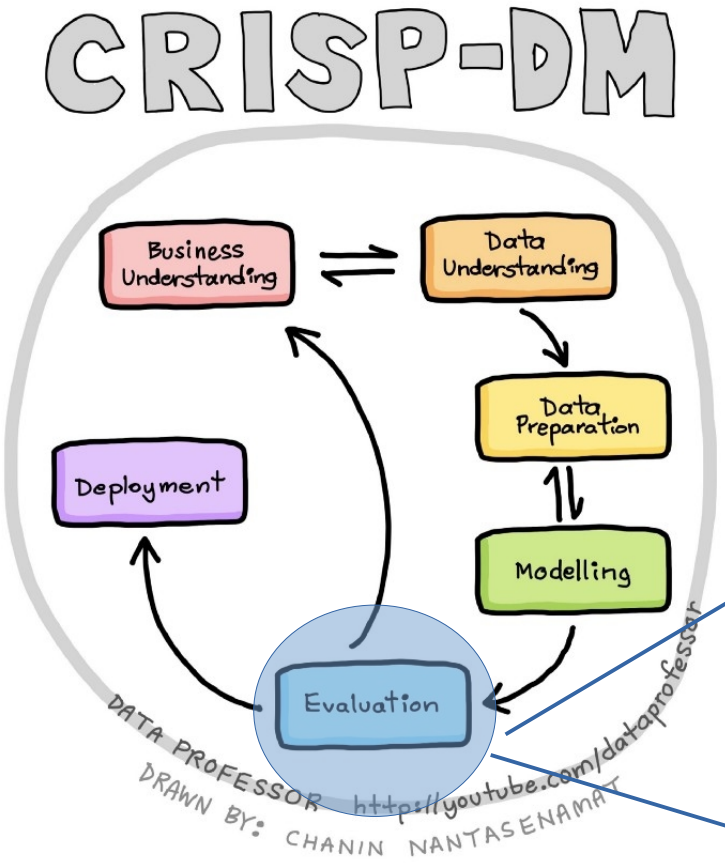
>

Questions?

THANKS!

Alejandro Fernandez Benages
Data Science Team

Evaluation



Evaluate the accrued results and review the process performed, to determine if the goals are met or not.

Rinse and repeat, some steps may need to be performed more than once, each time improving metrics and the understanding of the data.

Future, in this evaluation phase, some findings may ignite new project ideas for which to explore.

Is there any recommended **metric** that we can use to guide our project?

- minimum accuracy desired
- minimum % of rejected customers
- ..