Project Proposal

Credit One Banking Services

Understanding the Credit Limits and Lending of our Customers

<>

Process

STEP 1

Define the Goal

1

2

STEP 2

Collect and Manage Data

STEP 3

Evaluate and Critique the Model

3

4

STEP 4

Present Results and Document

STEP 5

Deploy and Maintain the Model

5

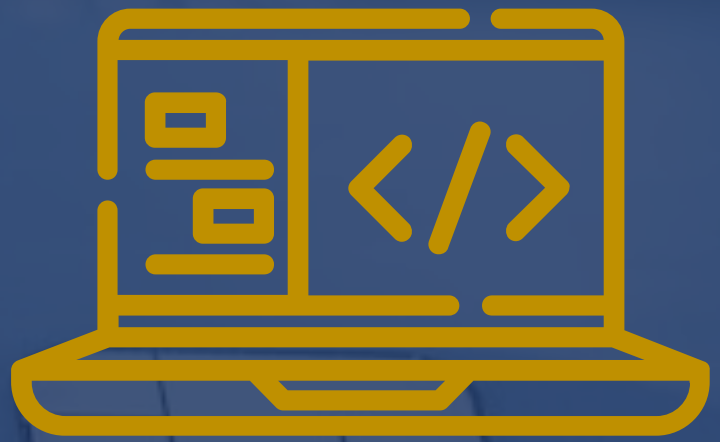
Process created by Zumel and Mount

Step One – Define the Goal

One Project, Two Goals

1. **Determine which potential customers are most likely to default**
2. **Determine how much credit balance to give customers**

Step 2 – Collect and Manage Data



All data comes from Credit One's external SQL server

Data Summary

30,000
customers

6 months
Billing and payment data

Age, Marital Status, Gender, Education
Demographic Data

Limit Balance, Default
Target Variables

Management

The data warehouse is kept in secure servers on our team’s private computers.

Data Issues

- There were a few rows whose values were what we wanted to name the columns, while the columns were merely X#. This was easily fixed.
- We assigned numbers to non-numeric values for easier parsing

	MyUnknownColumn	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
0	ID	LIMIT_BAL	SEX	EDUCATION	MARRIAGE	AGE	PAY_0	PAY_2	PAY_3	PAY_4	PAY_5	PAY_6
1	1	20000	female	university	1	24	2	2	-1	-1	-2	-2
2	2	120000	female	university	2	26	-1	2	0	0	0	2
3	3	90000	female	university	2	34	0	0	0	0	0	0
4	4	50000	female	university	1	37	0	0	0	0	0	0

	ID	Limit_bal	Gender	Education	Marriage	Age	Pay_1	Pay_2	Pay_3	Pay_4	Pay_5	Pay_6
1	1	20000	1	3	1	24	2	2	-1	-1	-2	-2
2	2	120000	1	3	2	26	-1	2	0	0	0	2
3	3	90000	1	3	2	34	0	0	0	0	0	0
4	4	50000	1	3	1	37	0	0	0	0	0	0
5	5	50000	2	3	1	57	-1	0	-1	0	0	0

Additional Useful Data

Credit Score

Income

Many credit card companies use credit score and income information to determine who to approve for credit cards and how high to set an individual's credit limit.

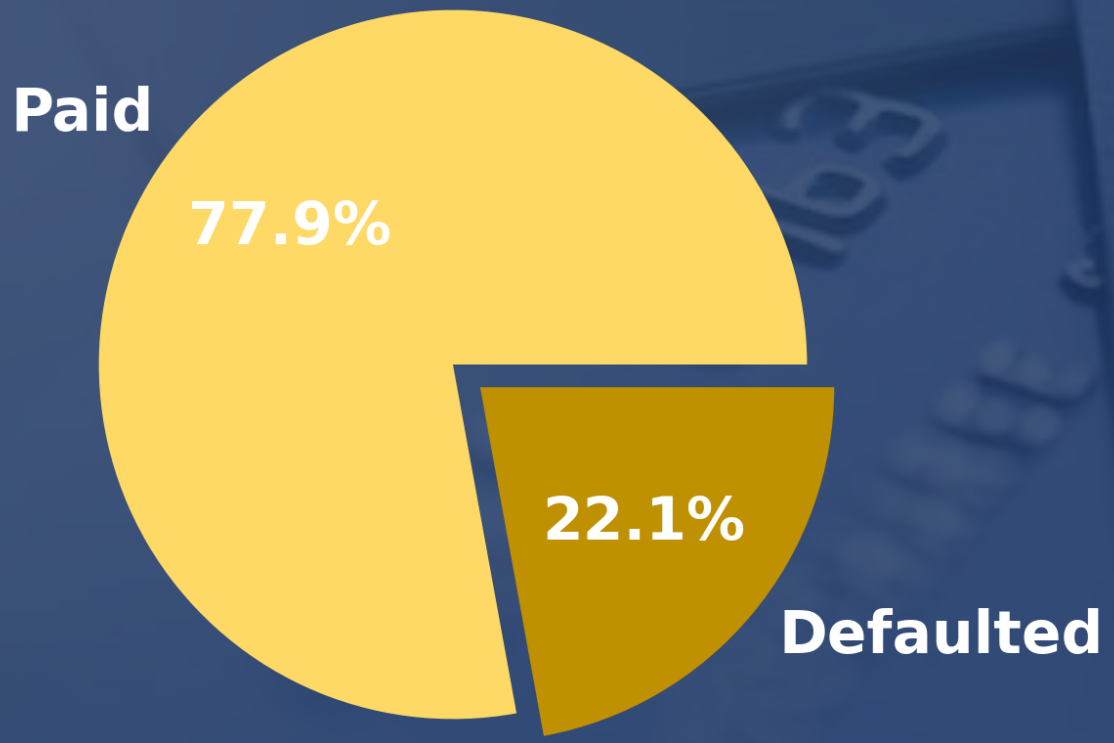
Step Three – Evaluate and Critique the Model

Initial Observations

- **The Limit Balance has a positive correlation with education, age, bill amount and pay amount, and a negative correlation with pay and default**

Defaulting has a positive correlation with pay and a negative correlation with limit balance

Defaults by Percentage



More than 1 in 5 creditors defaulted on their cards within the last 6 months

Step Four – Present Results and Document

- The most important aspect of the model is how it will reduce Credit One's spending on loans that are likely to default.
- When presenting, we should emphasize the reduction of their losses and include any other findings that will help their business.
- Finally, we will need to document the model for the team that will be running and maintaining it.

Step Five – Deploy and Maintain the Model

- Going forward, we can use the algorithm to vet new prospective Credit One members and review current ones
- As new data is gathered, the model we create may prove to be inaccurate or need updating.

STEP 1

Define the Goal

- Misinterpreting the goal can lead to a model that is not useful.
- Communication with management and careful monitoring of the model can prevent this problem.

1

STEP 2

Collect and Manage Data

- Some of the data may not be useful for new customers.
- There may be additional data that is helpful.
- If we cannot get more data, then we will have to make the most of the data we have.

2

STEP 3

Evaluate and Critique the Model

- We may have to try many iterations of the model to get the most accurate result.
- Proper preparation and analysis before modeling can mitigate this

3

STEP 4

Present Results and Document

- We may not be able to create an accurate model with the data
- While we may not always be able to accurately predict the target value, there may still be some useful information in the data.

4

STEP 5

Deploy and Maintain the Model

- As new data emerges, the model may prove to be less accurate than originally thought.
- We may need to revisit this problem when new data arises

5

Noting potential pitfalls that may occur along the way