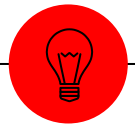
iFood CRM Data Analyst Case

Guilherme Gomes



Index

- Introduction
- Customer segmentation
- Predictive model
- Conclusion

1 Introduction

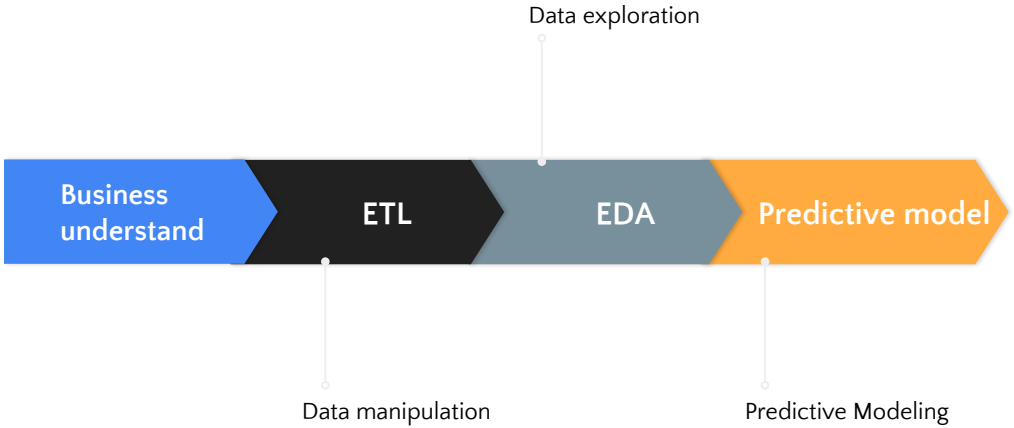
Introduction

After 3 years with a globally solid revenue, the company doesn't have a good growth profit perspective for the next 3 years. An alternative to solve it is improving the marketing campaign performance by using data analysis and machine learning.

In order to do this, we need to answer some questions:

- What's the main behaviour of the respondents?
- Is it possible to indicate how respondents will accept the campaign?

Analysis steps



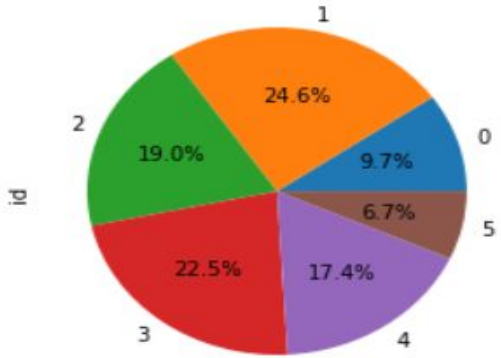
Tools:

- Python
- Pandas
- Scikit-learn
- Jupyter notebook

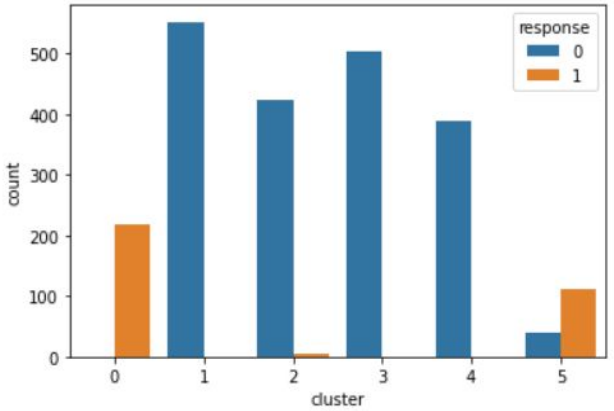
2 Customer segmentation

Customer segmentation

1. % of customers by cluster



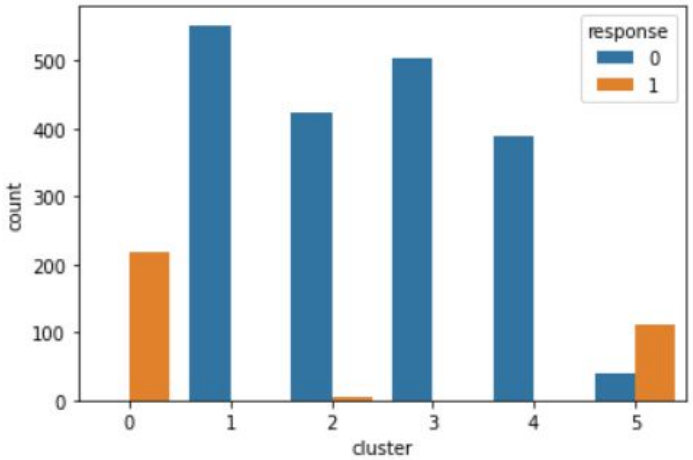
2. # of campaign success



- The K-Means method returned 6 customer segmentations (cluster).
- The second chart shows the people that buy a new item after the last campaign: being **0 the people that didn't buy** any item and **1 people that bought** some item.

Customer segmentation

2. # of Campaign Success



- It is possible to notice that **more than 300 persons** that are in **cluster 0 and 5** bought something after the campaign. Let's see the **main characteristics** of this clusters:
  - **Cluster 0:**
    - Income: $ 30,000.00
    - Retention time around 9 years
    - Age between 40 years and 60 years
    - Marital status is single
    - Education: Graduation
  - **Cluster 5:**
    - Income: $ 80,000.00
    - Retention time around 7 years
    - Age between 40 years and 60 years
    - Marital status is married
    - Education: Graduation

3 Predictive model

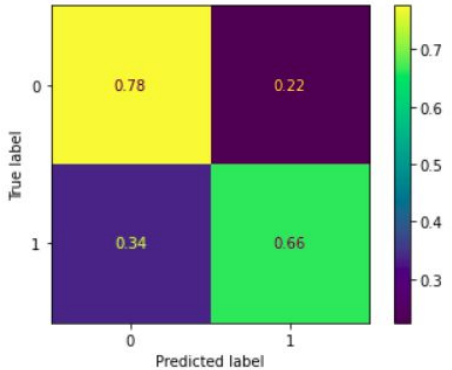
Predictive model

The model can provide a person with 76% of certain!

Model metrics

	precision	recall	f1-score	support
0	0.93	0.78	0.85	577
1	0.33	0.66	0.44	95
accuracy			0.76	672
macro avg	0.63	0.72	0.64	672
weighted avg	0.85	0.76	0.79	672

Confusion matrix



Precision: % of true positive

Recall: % of all true class

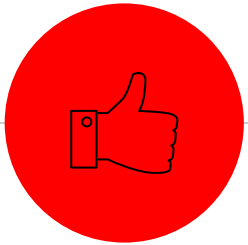
F1-score: harmonic average between precision and recall.

4 Conclusion

Conclusion

- With this model the success rate improved from 15% to 29%.

- The actual ROI is -45%, but we could improve it to 0,05%.



Thanks!

Questions?

Contacts:

- **Email:** gg.gomesguilherme@gmail.com
- **Linkedin**: https://www.linkedin.com/in/guilherme-gomes-06/
- **Github:** https://github.com/gomesgui06
- **Medium:** https://gomesguilherme.medium.com/