ORACLE
Consulting

PoliMI Challenge

A real case of Data Science problem

Oracle Consulting

Mar 8, 2022

«Pills» of Oracle Story

1977

1979
First Relational
DB in the Market

1986
IPO – NASDAQ

From 2004 ...
Acquisitions and
IT Market
consolidation

2009
Acquisition of SUN
MS ... Solaris,
SPARCand Java

2015
Oracle becoming
a Cloud Company

2018
Autonomous
Strategy

2022

Oracle becoming
Hyperscaler
Cloud

Our mission is to help people see data in new ways, discover insights, unlock endless possibilities.

Brand Awareness

Last year

60+ engaged business employees

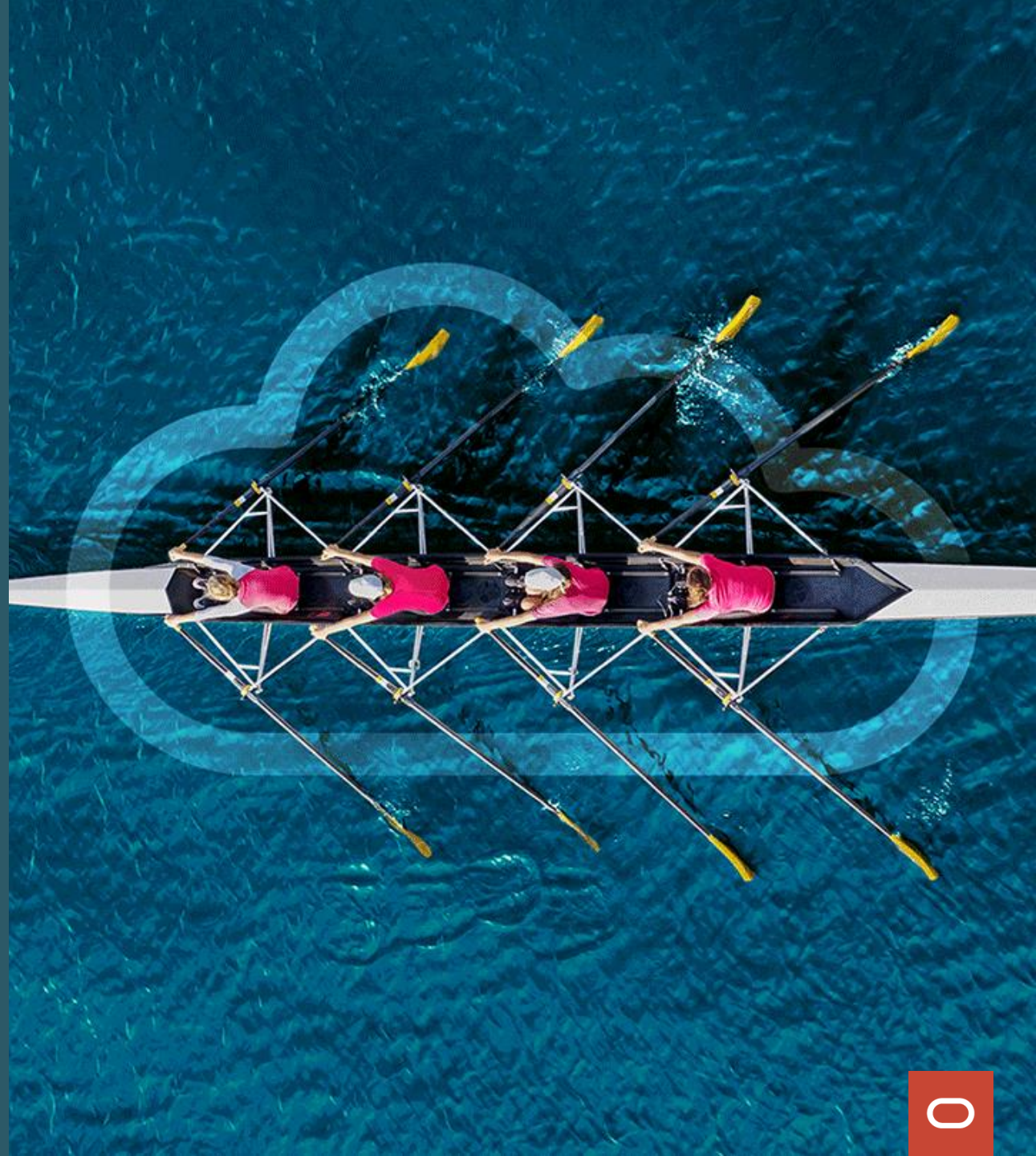
150+ hours of speeches

100 events (*company presentations, thematic workshops, recruiting days, mock interviews, roundtables…*)

Agenda

1. Use Case
2. Data Explanation
3. Unbalanced Classes Issue
4. Train & Test your Model
5. Classify New Observations
6. Features Importance
7. Next Steps to Improve your Incredible Work ☺

Use Case: Do you really know your customers?



Do you really know your customers?

Analyze consumers' behavior and preferences and show how well you can understand your customer base

Context and challenge

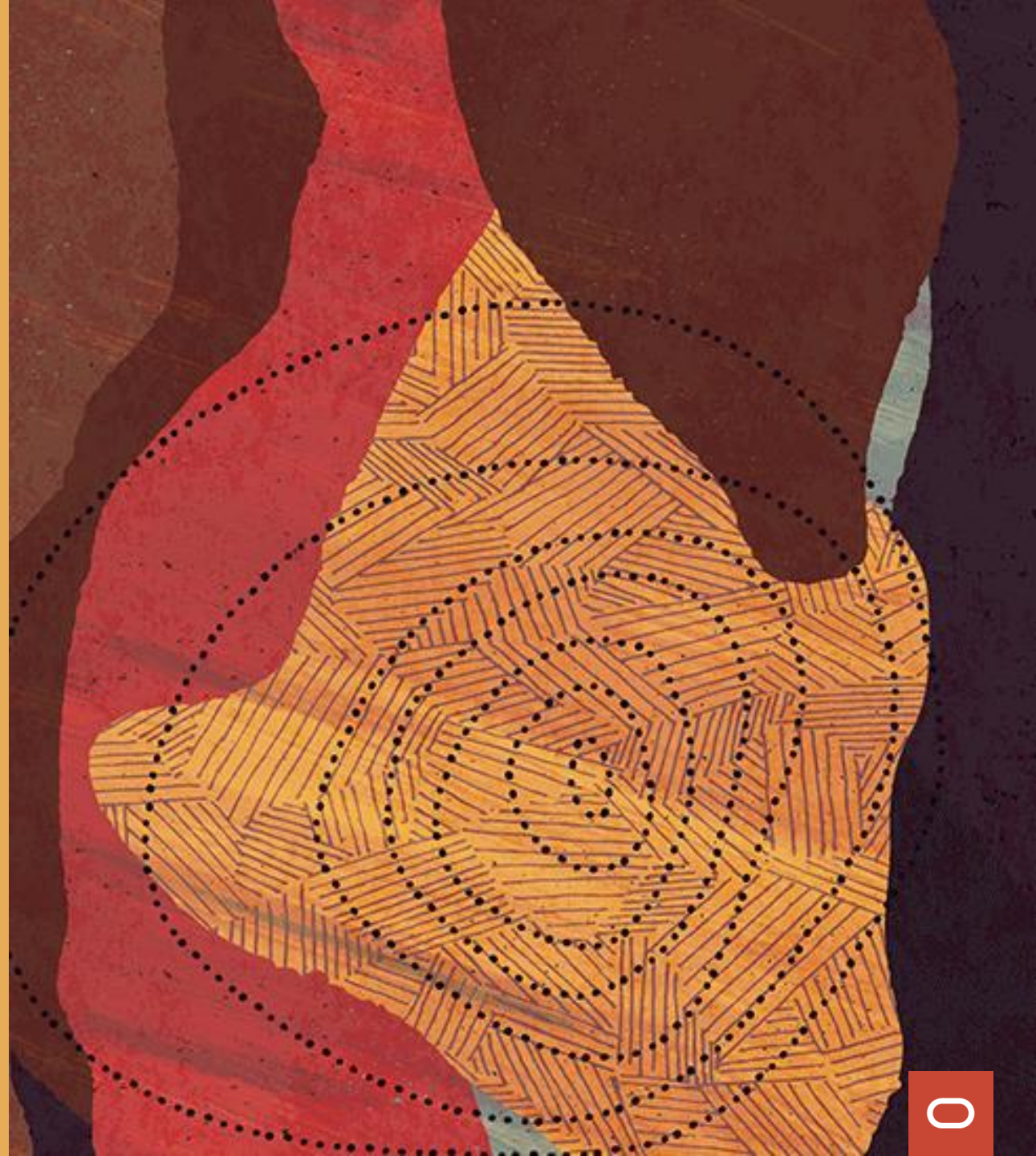
- A banking company needs to know if its customers are likely to buy a specific financial product based on their behavior and characteristics.
- The final aim is to suggest the best pool of customers to recommend a specific financial product and consequently increase their satisfaction.

Key objectives

- Apply a classification model on your target variable.
- Explain your classification outcomes trying to find the main features each customers' group.

Hands-On

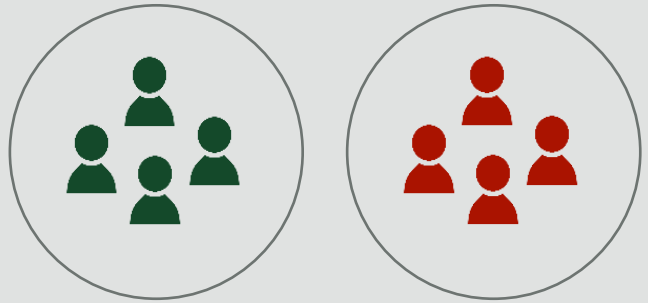
Let’s play!



Data Explanation

Entities involved

- **CUSTOMERS:**
  - Anagraphic Data
  - Job
  - Banking Balance
  - Default History
  - Loan History
  - Communication Details
  - Target Financial Product



Help us to identify wether a customer would be inclined to buy a target financial product

Unbalanced Classes Issue

How could you deal with the unbalanced nature of the dataset?

- Find the best technique to solve this problem (i.e. undersampling, oversampling, ...)

Train & Test your Model

You have a wide range of classification models. Enjoy training and testing one or more of them

- Show us the trained models (i.e. paramaters/hyperparameter set) and how you validated them
- Select the best one according to the F1 score metric

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

TP = True positive

TN = True negative

FP = False positive

FN = False negative

Classify New Observations

It's time to apply the best model to the new observations

- Use the optimal model to score the new observations (*test.csv*) and save the output *.csv*
- We know the outcome on these observations and will just compare them with your predictions.

This is the first step to rank you in the top positions ☺

Features Importance

Explain the classification results performing the features importance.

Next Steps to Improve your Incredible Work ☺

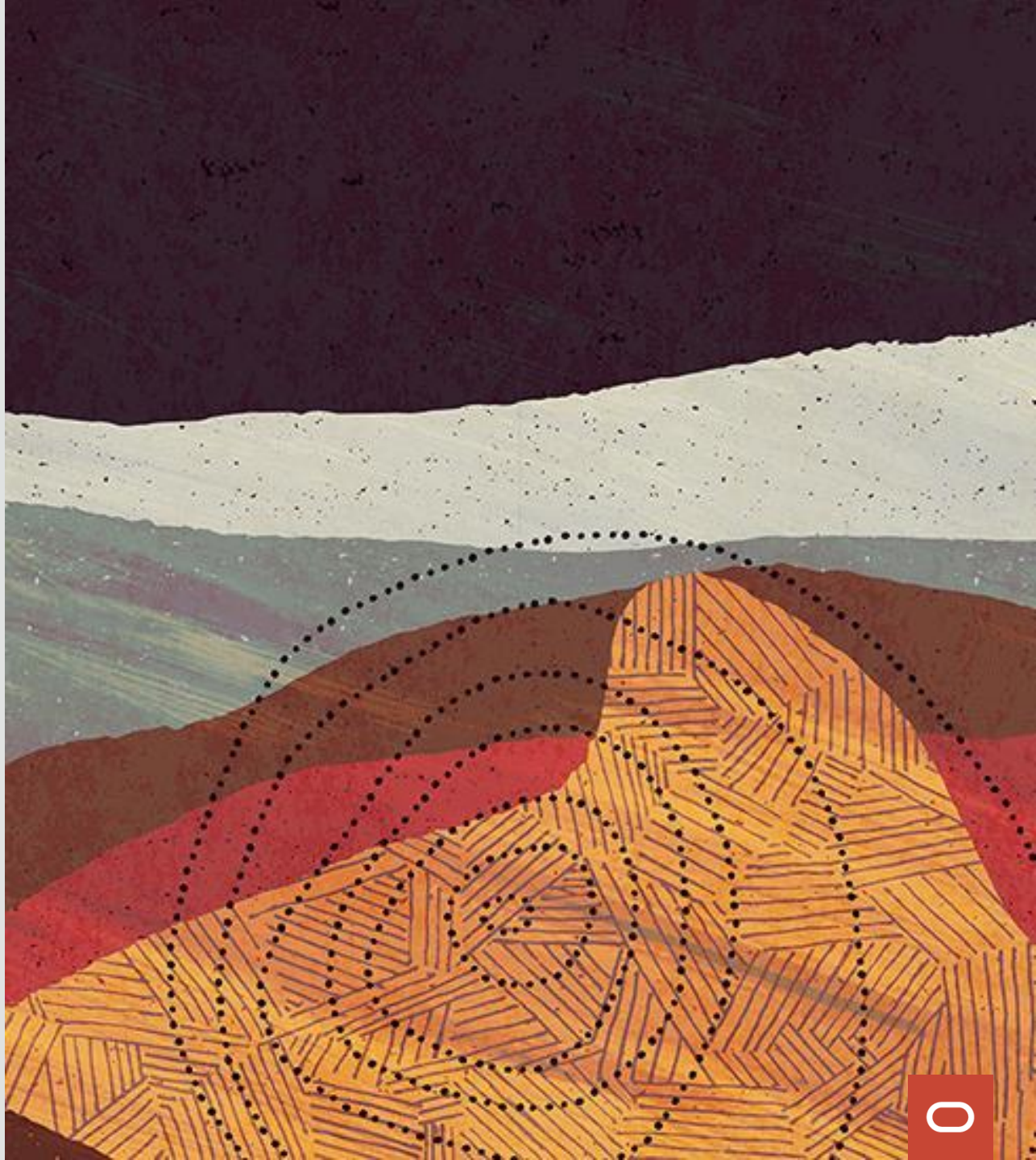
Give us advices to improve the work. We need new ideas and fresh minds

Resources

Download package (dataset + notebooks) + Submit your Solution :

https://tinyurl.com/OracleChallenge

Thank You and Good Luck




ORACLE