Data Science
Use Case Challenge

Raiffeisen BANK

OBJECTIVES

You have just been hired as a Data Scientist for a Telco Company

Congratulations!

First use case will be working on customer Payment Delays model.

- You are firstly expected to find out the reasons for customer churn by doing an Exploratory Data Analysis and suggest counter measures.
- Secondly you are expected to build a machine learning model and use it on new data to identify top 300 customers that are expected to delay in payments.

Raiffeisen BANK

Data Set Description

The data set was collected in the USA and includes 19 predictors:

state (categorical): The US state of the customer
account_length
area_code
international_plan (yes/no)
voice_mail_plan (yes/no)
number_vmail_messages
total_day_minutes
total_day_calls
total_day_charge
total_eve_minutes
total_eve_calls
total_eve_charge
total_night_minutes
total_night_calls
total_night_charge
total_intl_minutes
total_intl_calls
total_intl_charge
number_customer_service_calls.
The target variable is payment_delay (yes/no)

The Challenge

The overall goal is to determine if a customer is about to churn or not.

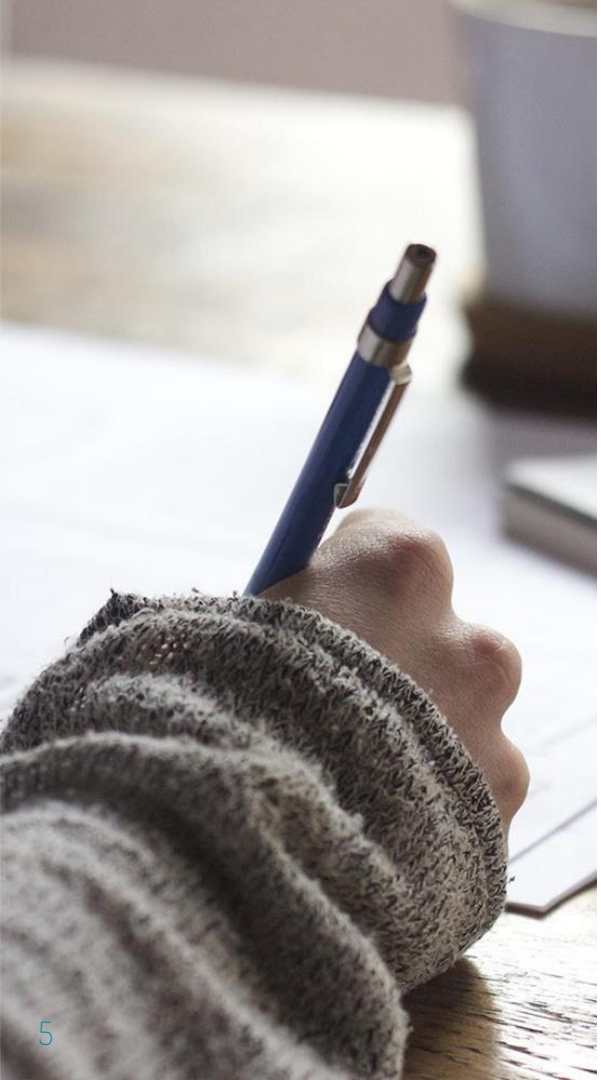
The following steps need to be accomplished:

1) Do an Explorative Data Analysis (EDA) and document interesting patterns, data issues and ideas.
Document your findings in an Jupyter Notebook file which can easily presented.

2) Find a model to predict if a customer will register a payment delay or not. Consider selecting your features and data pre-processing depending on your models used.

3) You are given a list (test.csv) of 2,000 customers to asses if they are likely to register payment delays in the following period. Since your action time is limited, you need to extract 300 customers from the list who will then be addressed directly by the company.
Please submit the indices [1,2000] of the first 300 customers from the test.csv list which are likely to register payment delays.

You will present your findings in a 10-15min talk addressing:
The results of the EDA.
Model used and performance characteristics.



Key Deliverables to be send before the interview

1) Jupyter Notebook with EDA and Modeling Approach
2) .csv file containing the indices of the first 300 customers that the Telco company needs to address (scoring the list from the test.csv data set)

Raiffeisen BANK

Thank you for the time! Good luck and we look forward meeting you!

Advanced Analytics – Risk Hub