Customer Churn Prediction using Machine Learning

Georgios Koliopoulos
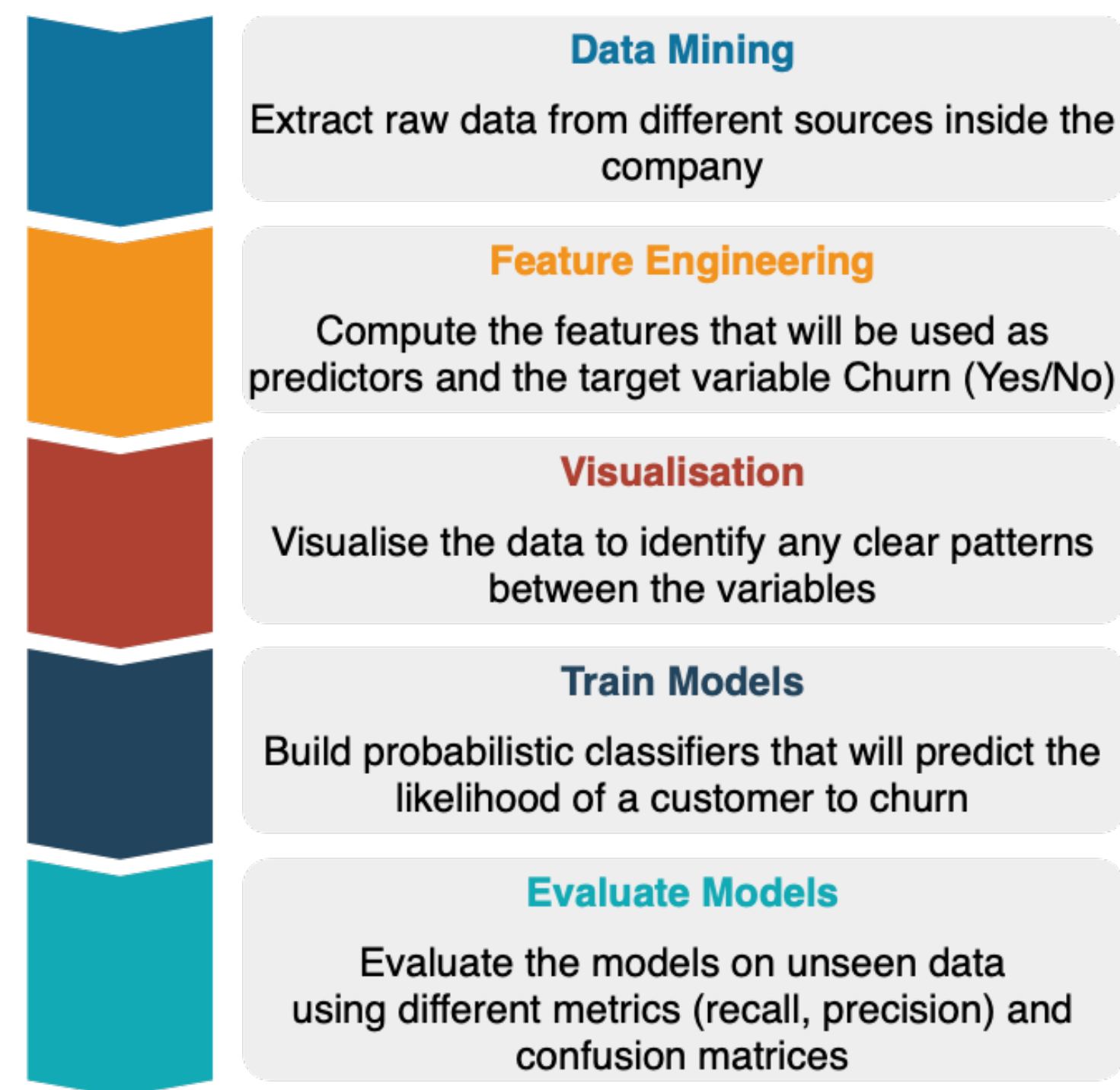**MSc in Big Data**

UNIVERSITY *of* **STIRLING**

Problem

Retaining customers costs less than acquiring new ones. The purpose of this project is to identify players of People's Postcode Lottery that are more likely to churn- cancel their subscription- the following month in order for the relevant departments to target those players and try to retain them.

Objectives

- Gain insights about the characteristics of customers that churn by finding and understanding any clear patterns.

- Create useful visualisation that would demonstrate the relationships between the data to the relevant departments and stakeholders.

- Build and evaluate different machine learning models (logistic regression, random forest, support vector machine, neural network) to find the most efficient one for the particular problem.

Process

Data Mining
Extract raw data from different sources inside the company

Feature Engineering
Compute the features that will be used as predictors and the target variable Churn (Yes/No)

Visualisation
Visualise the data to identify any clear patterns between the variables

Train Models
Build probabilistic classifiers that will predict the likelihood of a customer to churn

Evaluate Models
Evaluate the models on unseen data using different metrics (recall, precision) and confusion matrices

PEOPLE'S POSTCODE LOTTERY

python™   pandas

$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$

scikit learn