

# Model Selection Lab Sheet

## Preliminaries

The notebook 07 Model Selection Lab shows how to fit a  $k$ -NN model for the HotelRevHelpfulness dataset. It assesses three options:

- whether to use a StandardScaler, MinMaxScaler or no scaler.
- what  $k$  to use for  $k$ -NN
- what weighting policy

## Q1

Repeat this model fitting exercise to fit a Naive Bayes model for the same data. Consider the same scaling options and GaussianNB and BernoulliNB as classifier options.

## Q2

Find the best decision tree model for the HotelRevHelpfulness dataset considering max\_leaf\_nodes and the splitting criterion. The splitting criterion can be either 'gini' or 'entropy', you can select your own options for max\_leaf\_nodes.

## Q3

Which model would you recommend for this dataset?