### Assignment 1: Prototype [due 6 Feb 2024]

Build a **prototype** for*email spam classification*

1. in prepare.ipynb write the functions to
   * load the data from a given file path
   * preprocess the data (if needed)
   * split the data into train/validation/test
   * store the splits at train.csv/validation.csv/test.csv
2. in train.ipynb write the functions to
   * fit a model on train data
   * score a model on given data
   * evaluate the model predictions
   * validate the model
     1. fit on train
     2. score on train and validation
     3. evaluate on train and validation
     4. fine-tune using train and validation (if necessary)
   * score 3 benchmark models on test data and select the best one

Notes:

* You may download the email spam data from: <https://www.kaggle.com/datasets/jackksoncsie/spam-email-dataset>
* You may refer to <https://radimrehurek.com/data_science_python/> for building a prototype.
* You may refer to first 3 chapters of <https://www.statlearning.com/> for basic ML concepts.
* You may refer to the Solution Design example covered in the class as a guideline for the experiment design.