2. Classification:
   a. Download the data to predict Down syndrome in mice from
      https://archive.ics.uci.edu/ml/datasets/Mice+Protein+Expression#. The prediction problem is to
      either predict the genotype (binary) using the gene expression variables from DYRK1A_N to CaNA_N.
   b. Explore, visualize, and pre-process the data as appropriate, including developing a strategy to deal
      with missing variables. You can choose to impute the variable. The recommended way is to use
      multivariate feature imputation (https://scikit-learn.org/stable/modules/impute.html) [1]
   c. Train, validate varying at least one hyperparameter, and test at least two types of models: [2]
      i. Random forest
      ii. Support vector classification using RBF kernel
      iii. Neural network with single hidden layer (output layer should be have softmax activation)
   d. See if removing some features systematically will improve your models using recursive feature
      elimination https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFECV.html). [1]