Exercise 1: Hypothesis Space, Capacity, Regularization

(a) Simulate a data set with $n = 100$ observations based on the relationship $Y = \sin(x_1) + \varepsilon$ with noise term ε following some distribution. Simulate $p = 100$ additional covariates $x_2, \ldots, x_{101}$ that are not related to Y.

(b) On this data set, use different models (and software packages) of your choice to demonstrate

- overfitting and underfitting;
- $L1$, $L2$ and elastic net regularization;
- the underdetermined problem;
- the bias-variance trade-off;
- early stopping (use a simple neural network as in Exercise 2).