The aim of this exercise: Compare prediction properties of a normal classification tree with ensemble methods. You can use pre-installed packages such as **caret** or **xgboost** for boosting.

Exercise 1:

Consider the following data generating process in which we have $n = 500$ observations and P covariates $X_j \sim \mathcal{N}(0, \sigma_j^2)$ for $j = 1, ...P$ and covariance matrix Σ. y_i is generated by some nonlinear function of $\mathbf{X}$ of your choice, this should include some (possibly higher order) interactions.

- Generate the data according to the dgp described above and fit a classification tree with optimal pruning and a boosted tree.
- Compare the test classification errors for both methods.

Exercise 2 (Simulation Study)**:**

The goal here is to think about how a regression tree makes its predictions and how boosting improves these properties.

- Propose a dgp that will improve the boosting classification error vs. the traditional regression tree. Illustrate the properties of your DGP in some simplified graphs.
- Propose and implement another ensemble method of your choice (either bagging or random forests) and benchmark these methods against boosting within the frame work above.
- How would you quantify the impact of a single covariate on the quantity of interest? Can this be used as a "diagnostic tool" to determine whether we should use bagging, boosting or random forests? Explain you answer.