**Variance-bias tradeoff exercise**

**José Luis Barreiro Tomé**

In this additional exercise to the course FYS-STK4155 I performed an analysis of the bias-variance tradeoff using three of the methods discussed in the course: Linear Regression (OLS, Ridge and Lasso), deep learning (feed forward neural network), Ensemble method (random forest) and support vector machines.

The analysis is performed on a classification problem, as a function of the complexity of the model. Here I have used the Franke function used in the Projects 1 and 2 of the course.

The bias-variance tradeoff:

The error for any supervised Machine Learning algorithm can be decomposed in Bias and Variance. There is a tradeoff between the model’s ability to minimize these two components. Bias indicates the difference between the model’s predictions and the true values. A model with high bias oversimplifies the prediction and underfit the data. The variance deals with the spread of the data. A model with high variance fit the training data closely but it’s not able to fit new data which hasn’t seen before. Such models are said to overfit the data. Finding the right balance between both components is called the bias-variance tradeoff.