Elements of Machine Learning
Assignment 1

William LaCroix (7038732)
wila00001@stud.uni-saarland.de

Problem 2 (T, 6 Points). **Bias-Variance Trade-off**

1. [5pts] Prove that the expected test mean square error (MSE), for a given value x_0, can always be decomposed into the sum of three fundamental quantities: the variance of $\hat{f}(x_0)$, the squared bias of $\hat{f}(x_0)$ and the variance of the error terms ϵ.

Let us begin with the definition of the average, or *expected* value of the squared difference between the predicted ($\hat{Y}$) and actual value of Y: $E(Y - \hat{Y})^2$. Given an estimated function, $\hat{f}$, and a set of predictors, X, we may predict $\hat{Y} = \hat{f}(X)$, thus $E(f(X) - \hat{f}(X))^2$. ISLR 2.1 tells us that, if X and Y are fixed, the only error comes from the irreducible error $\epsilon : E(f(X) + \epsilon - \hat{f}(X))^2$, rewritten in 2.3 as the sum of the irreducible error $Var(\epsilon)$, where 'variance' is the average of the square of the term, σ in this case, so $E[(X - \mu)^2] = AVE(\sigma^2) = Var(\epsilon)$, and the reducible error $[f(X) - \hat{f}(X)]^2$, which evaluated at x_0 gives us $E[(y_0 - \hat{f}(x_0))^2] + Var(\epsilon)$

Through neutral element addition, polynomial expansion, and reduction of terms (graciously swept under the carpet in ESL 2.25) the first half of 2.25 lets us rewrite as:
$E[\hat{f}(x_0) - E(\hat{f}(x_0)]^2 + [E(\hat{f}(x_0) - f(x_0)]^2 + Var(\epsilon)$, satisfying the first equality.

As with $Var(\epsilon)$, the variance of $\hat{f}(x_0)$ is the average squared difference of $\hat{f}(x_0)$ minus the mean $\hat{f}(x_0)$, giving us $E[\hat{f}(x_0) - E(\hat{f}(x_0)]^2 = Var(\hat{f}(x_0))$, and since bias is defined as the average difference between the true value, and our model's predicted value, we rewrite $[E(\hat{f}(x_0) - f(x_0)]^2$ as simply $[Bias(\hat{f}(x_0))]^2$, giving us $E[(y_0 - \hat{f}(x_0))^2] = Var(\hat{f}(x_0)) + [Bias(\hat{f}(x_0))]^2 + Var(\epsilon)$

□

2. [1pts] Explain in your own words irreducible and reducible error as well as their difference.

Irreducible error is error due to the somewhat random nature of the phenomenon being observed, or in the measurements being taken. Reducible error is error due to the inexact fit of the estimated function when compared to the true underlying function. Reducible error can be reduced by modeling a function more closely to the actual underlying mechanism, whereas irreducible error cannot be managed or mitigated using statistical methods.