**Christopher Tran**

**Problem 1 (2 Points)** Suppose that the data for study is where is a p-dimensional prediction vector. In regression analysis, we assume that the model has the form  . (Hint: Page 314)

**Problem 2 (2 Points)** Suppose that the data for study is where is a p-dimensional prediction vector. In regression tree analysis, we assume that the model has the form  . (Hint: Page 314)

**Problem 3 (2 Points) (True/False)** Even if the true function form of the model is well approximately by a linear model, the model build using decision trees is still better than the model built using regression.

**Problem 4 (10 Points)** Based on the textbook, Decision Trees for either classification or regression have four advantages over the traditional regression and logistic regression methods: **▲Trees are very easy to explain to people. In fact, they are even easier**

**to explain than linear regression!**

**▲ Some people believe that decision trees more closely mirror human**

**decision-making than do the regression and classification approaches**

**seen in previous chapters.**

**▲ Trees can be displayed graphically, and are easily interpreted even by**

**a non-expert (especially if they are small).**

**▲ Trees can easily handle qualitative predictors without the need to**

**create dummy variables.**

**▼ Unfortunately, trees generally do not have the same level of predictive**

**accuracy as some of the other regression and classification approaches**

**seen in this book.**

**Problem 5 (2 Points)** Suppose that be a random sample from a normal population with mean and variance . The variance of the mean is σ^2/n. . This means that we can reduce the variance of almost all statistics through averaging.

**PART 2 AND PART 3 CAN BE FOUND AT** [**https://github.com/ctran301/DM/tree/master/ASS06**](https://github.com/ctran301/DM/tree/master/ASS06)