Week 11

MATH 4400

 Jaden Jackson, Weilun Mai, Nicholas Smith

Clearview Live Research

Clearview Live is a company based out of Roy, Utah. They create call-center analytics software for any-type of call center, anywhere. Possibly they have international clients. The nature of the data then is that we know it is call center data. Because the company is based in Roy, Utah, perhaps we can just assume a MST timezone for simplicity/compatibility with our local machines. Unfortunately, because they work with any type of call center, we do not know much about the nature of the calls. Most of the information about the data will need to be interred directly from the data itself.

Model Approaches

Because both models (call volume and AHT) have a numeric response, we will need to select a supervised regression approach.   If we go with some sort of linear regression approach, we will likely try a ridge regression (if we keep our number of features small) otherwise a lasso (if we separate out a lot more categorical features). Going with the linear approach, we will need to be careful with our predictor selection, making sure to apply good transformations or appropriate categories to each one.

Otherwise, we will go with some sort of decision tree regression, such as a random forest or boosting. This direction will also need us to be careful with how we select and apply each feature.

PCA Seminar

What stood out to me about the PCA lecture is that PCA makes components that are completely linearly independent. Also, a subset of these components can be used to explain most of the response variance with a linear regression. Therefore, PCA is a good way to go from let's say 20 predictors, down to just several principal components that still explains most of the response variance. What also stood out to me is that the components are sometimes very difficult to interpret; however, sometimes an intuitive understanding can be gleaned from the values of the entries. For example, the speaker called one of his components "maturity" because the "age" coordinate was high, and all the other entries were low. Finally, I also liked the tip he gave that most the time the components that have a corresponding eigenvalue of less than 1 can be discarded as not significant enough.

Goldman Sachs seminar

What stood out to me about the Goldman Sach's lecture, is that a well formatted, targeted resume is very important to the job search. Also, it is important not to get too disappointed if it is taking a while to hear back from the company, because every company will operate on different time scales. I liked his advise that in real-world problems, there is no answer key. Therefore, the most important skill we can have is attention to detail, a healthy skepticism about if we are using the right data and drawing the correction conclusions. He also stressed the skills of being adaptable to learning new applications and tools, and that communication is key. Being able to distill a complex analysis in a concise, insightful narrative is very valuable.