**Andres Forero**

**Data Science Intensive**

**Potential Capstone Ideas**:

1. **Predict Unbilled Accounts Receivable balance:**

The Unbilled Accounts Receivable is an account that I tried to predict using exponential smoothing at work for a cash flow forecasting project. The Root Mean Squared error was very large (in the $millions of dollars), and thus my model was not too practical in terms of predicting the balance within a reasonable range. I was not able to find good features to predict the account balance.

For this reason, this could be a good Capstone project to try out, given I could use Principal Components analysis to help me to narrow down the most important features, and thus reduce the RMSE to a reasonable amount.

1. **Predict Win / Loss (binary) of Sales pipeline opportunities (leads):**

At work we use the Salesforce system to manage leads that enter the sales pipeline. The leads go through a Qualification stage whereby a decision has to be made whether to Bid or Not Bid. The leads that get bid on, progress to the Proposal / Price Quote stage, then the Review / Negotiation stage. There is a period of time whereby the lead is “Outstanding”, meaning the client is deciding whether to give the contract to my company (Win / Lose decision). I would like to apply a categorization machine learning algorithm to try to predict the opportunities my company will Win or Lose. This will be useful because it allows us to predict the future revenue that will be generated from our sales pipeline opportunities.

1. Use Zillow API data to predict house prices around my neighborhood in Raleigh.

Exactly how the title describes it. Use room size, square footage, school district, neighborhood and any other features that I am able to obtain from Zillow to predict the $ price of houses in my neighborhood.