Trevor McGlaflin

March 18, 2022

STAT 287

Project Check-in #1

**Working Title:** What Factors Drive Employee Turnover?

**Summary**

The goal of our project is to figure out which factors about employment lead to a higher employee turnover. Our data set contains fields such as salary, time since last promotion, time with same manager, stock option amount, and other benefits. Then, the same people were asked later if they stayed at that company. This field will be used as our dependent variable in our machine learning models.

**Major Activities Planned for This Week**

1. See if we can find a better data set
2. Start doing some EDA with the data to uncover basic relationships
3. Research which ML algorithms may be good for our analysis

**Major Activities Accomplished This Week**

1. Start doing some EDA with the data to uncover basic relationships
2. Research which ML algorithms may be good for our analysis

**Open Challenges and Questions**

1. **Our data set currently only has 1400 records in it. Is that enough records to draw any significant conclusions or would you recommend finding a larger data set?**

**Major Changes to Research Plan**

* None this week

**Revised Week by Week Timeline**

**Week of Mar 14**

1. See if we can find a better data set
2. Start doing some EDA with the data to uncover basic relationships
3. Research which ML algorithms may be good for our analysis

**Week of Mar 21**

1. Create a pipeline that cleans the data and adds any calculated columns
2. Start working on some plots

**Week of Mar 28**

1. Finalize 5 or 6 good-looking plots that show major trends

**Week of Apr 4**

1. **Start experimenting with machine learning algorithms**

**Week of Apr 11**

1. **Finalize one or two machine learning algorithms to include in our final analysis**
2. **Include a comparison of the accuracy of these algorithms**

**Week of Apr 18**

1. **Prepare for presentation**