Joshua Harkness

STA6704

Prof. Smith

<https://github.com/darkhark/2019JAX_PlayByPlay> under data and Assignment02

1. A screenshot of a cell phone

   Description automatically generatedThis was done throughout two scripts in the program. The dummy variables are created at the bottom of obtainPassData.R, but some were removed in the DimensionReduction.R script
2. PCA
3. T-SNE

A screenshot of a social media post

Description automatically generated

1. NMF - Basis

A screenshot of a cell phone

Description automatically generated

Using the scatterplots alone, there is not a lot of information to gain based on the y value created. I created the y value myself where I determined the quality of the play based on turnovers, touchdowns, yards given up, and a few combinations of those variables. My end goal was to determine if the given variables could help predict the where the defense was weak. One example could include on third down, late in the game, the defense appears weak when the opponent throws to the left side of the field. PCA’s scatter plot looks like a giant glob, but the graph produced ggfortify did seem to provide some insight into possible hidden details within the data. The image can be seen below. Although the image appears to be a blob, the plays do seem to range in quality from left to right, although I can’t 0, the worst quality of play. T-SNE appeared to create a few loosely grouped clusters, but nothing that appeared concrete and it certainly did not align with the quality of the play. Lastly, the nonnegative matrix factorization appeared to create 5 unique clusters, although none of them appear unique to the quality of play. I think it’d be worth while to run a cluster analysis on the NMF data. While the quality of play does not appear dependent on any of these dimension reduction techniques, there does appear to be some clear grouping in the NMF scatterplot. For my y variable, I would say PCA was the best, but for possible clear clusters on an unknown y, I would choose NMF.

A close up of a map

Description automatically generated