Jake Arendsen, Adil A Kumar, Zihan Wang, Ivy Zhou

ISOM 671 Project Proposal

4 February 2020

FIFA vs. Real World Comparison Project

1. We want to take advantage of big data tools to do a thorough analysis/comparison of a soccer player’s real-world performance compared to how they are statistically rated in FIFA 2019, EA’s best-selling soccer videogame.
2. We’ve found a Kaggle dataset that has ratings for 18,000+ professional soccer players (<https://www.kaggle.com/karangadiya/fifa19>) in the FIFA game, and we can take advantage of the web scraping tools we learned in Sports Analytics to pull data for their professional performance.
3. Among other things, we want to compare statistical ratings of players in FIFA versus how they perform in real life in those areas. Also, we want to attempt to use a player’s real-world statistics to predict their FIFA rating. Finally, we’re considering attempting to cluster players based on their FIFA ratings to see if there’s any “typical” players.
4. We will use Spark MLLib for exploratory data analysis, data preparation, potential dimension reduction, and modeling. In case we use iterative methods, we will use Spark ML as well.