Filipp Krasovsky University of San Diego MSADS 501: Foundations of Data Science Section 2 1-24-2021

Journal Entry 1: Strategion

The degree to which society trains an algorithm by providing data is dependent on the data's degree of separation from the individual. Wind turbine performance is unlikely to incorporate social bias into the modeling process compared to job candidate evaluation, which factors in a variety of intrinsically human traits. In doing so, the latter forces us to categorize and rank some demographics over others based on their aggregate historical performance. In the case of Strategion, able-bodied athletes are given preeminence over disabled employees through overrepresentation in the labor pool. Notably, Strategion did not intend for this to be the case – including athletic activity in the analytics product created a proxy by which disabled people were likely to be excluded from the hiring process. In these instances, dimension reduction is a useful way to minimize bias – in the case of Strategeion, this means removing proxy dimensions such as athletic activity which act in tandem to disability status and being cognizant of other dimensions that may exclude protected groups in the ABT. With regards to data combating negative biases in society, analytics can provide illumination about the falsehood of stereotypes – for instance, a study showing an ethnic group's expressed interest in obtaining a bachelor's degree over the course of their lifetime (as opposed to college admissions for high school seniors only) might be able dispel a myth that a certain ethnic group is "lazy" or "uneducated". These falsehoods are sometimes propagated using statistics that are intentionally deceptive - thus, broadening the time-series over which a given individual pursues a postsecondary education can be a useful counterbalance.