Geneway Tang

MI-462

Dr. Pursell

5 September 2022

Steps I can take to reduce bias in ML and AI systems

Watching the Gebru-Denton tutorial and listening to the perspectives they had on working with data in machine learning and artificial intelligence has opened my eyes to the risks of bias and discrimination. We must be mindful when using datasets within our projects and take steps to reduce bias and discrimination in our work.

In the first part of the tutorial, one thing that Gebru mentioned that was very striking to me was the issue that arose from poor training datasets for facial recognition models. She mentioned that most of the training data had included only Western faces, which would cause poor recognition in darker skin tones, yielding very high error rates in the model. This was caused by a significant underrepresentation of darker skin tones for facial analysis. This potentially harms marginalized communities even more from the discrimination that they already face daily. Additionally, not only image recognition models, text models with data taken from platforms such as Reddit and Twitter are often damaging to marginalized populations due to the unwelcoming nature of the sites to these communities.

Denton's tutorial spoke on the inescapability of bias in datasets and how there is no universal understanding of many concepts that we may be striving for. A very interesting example that she gave was the lobster and trout classification where they were linked with food and trophies, respectively. She states that there is always a particular view of the world embedded within data and while there is always bias in datasets, not all are created equal. She

states how not bias is not necessarily unmeaningful and we need to recognize that the goal isn't to get a completely unbiased dataset but recognize the particularities of a dataset, and what particularities should be embedded.

I believe that Denton's statement on recognizing the particularities of a dataset is extremely important and I will be practicing it within my data projects. For instance, if I am training a model that will help predict diseases in different individuals, I will make sure to include different demographics and regions where there could be differences could arise, to minimize bias within my model. Another point that Denton mentioned that I think is very important to incorporate into my work is the importance of socially responsible data practices. It is important to use data that is well maintained and well documented to ensure high quality of data. There seems to be a large disparity between model work and data work, where there is a huge lack of investment in dataset maintenance. In my work, I will strive to seek datasets from original sources and not derivative datasets, and ensure that it is up to date, so I can mitigate issues related to dataset quality.

Upon completion of the Gebru and Denton tutorials, I recognize that while it is nearly impossible to escape bias within datasets, I can still keep in mind particularities within a dataset in the context of the model I am trying to create.