Personal AI Ethics Statement

|  |  |  |
| --- | --- | --- |
| **Principle** | **What I will do** | **What I will not do** |
| Protection of Youth and Minors | Ensure datasets were gathered without taking advantage of children. | Participate in a business plan that attacks the innocence of children for monetary gain. |
| Transparency | Record my actions while building algorithms with datasets and make public the intent of my analysis. | Hide the actions and intent I take when analyzing a dataset or constructing an algorithm. |
| Dataset Bias | Research and educate myself on the inherent bias within the dataset. Ensure that the dataset is representative of the population my company is trying to reach, while not excluding people. | Trust that the dataset I have obtained completely bias free and should be used without any hesitation. |
| Algorithmic Bias | Determine what values are present in the construction of my algorithm. Determine if these biases will harm the analysis I will conduct, or people whom the analysis is about. | Assume there is no bias within the algorithm that is of note. |
| Automating Ethics | Never construct an AI algorithm assuming that it will operate within an ethical guideline. | Assume there are no ethical issues present within an algorithm or analysis. |
| AI Discrimination | Ensure the dataset that the algorithm is trained on does not have inherent discriminatory biases. Also determine if the algorithm is making assumptions about certain demographic factors (sex, race, etc.), as this has been shown to be an ability of ML. | Assume that the training dataset if rid of discriminatory biases. Sit back and not analyze the ethical implications of my analysis. |
| Personal Data Privacy | Ensure the personal data of my company’s clients will be protected and not made available to the public or to 3rd party organizations who seek to create potentially harmful algorithms. | Not care about the protection of my company’s client’s personal data. Treat this issue as a problem that is not mine to solve. |
| Evil Actor Problem | When creating an algorithm, I will ponder if it can be used to harm certain groups of people. | Not think about the consequences of my algorithm falling into the wrong hands and being used for nefarious purposes. |
| Environmental Effect | Realize the analysis I conduct, or the algorithms I develop, can have an effect, both positive and negative, on the surrounding environment. I will weigh the positive and negatives of such algorithms and determine if there will be a harmful environmental effect. Such an example would be the harmful rhetoric of Microsoft Tay’s chatbot. | Stay ignorant on the potential harmful effects I can create on society with my work. |