Fair Decision Making

David Yang, Selena She, Amy Feng, Xander Goslin

Summer 2023

1 Background Reading

1.1 Generative Adversarial Networks

Definition 1 (Generator)**.** *A* ***generator*** *is a neural network that learns to generate realistic samples by transforming random noise into data samples that resemble the training data.*

Definition 2 (Discriminator)**.** *A* ***discriminator*** *is a neural network that aims to distinguish between real and fake examples.*

More formally, a generative model captures the data distribution and a discriminative model estimates the probability that a sample came from the training data rather than G. "The generative model can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection, while the discriminative model is analogous to the police, trying to detect the counterfeit currency."

Definition 3. ***Adversarial nets*** *refer to the specific case where the generative model generates samples by passing random noise through a multilayer perceptron, and the discriminative model is also a multilayer perceptron.*

1.2 Denoising Diffusion Probabilistic Model / Denoising Diffusion Implicit Model