**Graded Discussion 4**

Bias is a significant concern when training machine learning models. In a technical sense, bias is the error arising from capacity limitations of the machine learning model. In other words, the limited ability of a model to fit the training data.

An important, and often overlooked or ignored, source of bias is the limitations in the training data themselves. A common source of such bias is having an important class or category that is significantly under-represented in the training data. These situations occur in applications ranging from social media, autonomous vehicles to fairly mundane industrial applications. The results can range from embarrassing (my dog is not a skunk!) to dangerous (autonomous vehicle not recognizing a child).

Think about some case you have either recently encountered or read about where biased training data has led to dangerous or unethical behavior by a deep neural network agent. What problem did you perceive? What steps can you think of that might mitigate this problem?

Make a post of about 100-300 words here in Piazza. Then post a comment of suggestion on another students post of about 100-200 words.  Posts are due by May 10. Do not post respond in Canvas.