Works Cited

All of the resources cited in Course 2 Week 2, in one place. You are encouraged to explore these papers/sites if they interest you, especially because this is an important topic to understand. They are listed in the order they appear in the lessons.

From the videos:

- Hyperspherical Variational Auto-Encoders (Davidson, Falorsi, De Cao, Kipf, and Tomczak, 2018): https://arxiv.org/abs/1804.00891

- Generating Diverse High-Fidelity Images with VQ-VAE-2 (Razavi, van den Oord, and Vinyals, 2019): https://arxiv.org/abs/1906.00446

- Conditional Image Generation with PixelCNN Decoders (van den Oord et al., 2016): https://arxiv.org/abs/1606.05328

- Glow: Better Reversible Generative Models (Dhariwal and Kingma, 2018): https://openai.com/blog/glow/

- Machine Bias (Angwin, Larson, Mattu, and Kirchner, 2016): https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

- Fairness Definitions Explained (Verma and Rubin, 2018): https://fairware.cs.umass.edu/papers/Verma.pdf

- Does Object Recognition Work for Everyone? (DeVries, Misra, Wang, and van der Maaten, 2019): https://arxiv.org/abs/1906.02659

- PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models (Menon, Damian, Hu, Ravi, and Rudin, 2020): https://arxiv.org/abs/2003.03808

- What a machine learning tool that turns Obama white can (and can't) tell us about AI bias (Vincent, 2020): https://www.theverge.com/21298762/face-depixelizer-ai-machine-learning-tool-pulse-stylegan-obama-bias

From the notebook:

- Mitigating Unwanted Biases with Adversarial Learning (Zhang, Lemoine, and Mitchell, 2018): https://m-mitchell.com/papers/Adversarial_Bias_Mitigation.pdf