Title: Coded Bias
Group Name: City gorls
Members: Elizabeth Breslin, Grace Berrien, and Lilleth Snavely
Leader: Elizabeth Breslin

Narrative:

In recent years the use of artificial intelligence for facial recognition by law enforcement has become a more frequent practice. However, recent studies have found that the facial recognition algorithms exhibit racial bias and non-white faces are incorrectly identified more often than white faces. One study by the National Institute of Standards and Technology found that "facial recognition technologies falsely identified Black and Asian faces 10 to 100 times more often than they did white faces." [1]. In another facial recognition tool analysis that is used by law enforcement was found to have misdentified 35% of dark skinned women while only 0.8% of light skinned men were misidentified [2].

While often the assumption in using automation instead of humans is that there is no bias and all decisions are completely objective, algorithms such as facial recognition technologies provide evidence that this is not the case. Technology can inherit the bias of whoever created the technology. Since 80% of all AI professors are men [2], the lack of diversity within the development phase of technology will be implemented in the results as well. Put more clearly, due to the majority of white faces being used to train the facial recognition models, the technologies work best on white faces. Our group is curious to see if diversifying the training data to be composed of a more even distribution of racial groups will result in more accuracy in classification of non-white faces.

Hypothesis:

If we train a model to classify face images as either male or female with only white faces as opposed to all racial groups, it will be at least 10% less accurate in classifaction of males and females for all racial groups in comparison to a model trained with images of all racial groups.

Modeling Approach:

In order to classify our facial images as either male or female, we will use a logistic regression model. A logistic regression model will be used due to its reduced computational time and its lack of complexity. This less sophisticated model will be more feasible to build and train within our time and resource constraints. A logistic regression model should be sufficient for this analysis due to the fact that the classification being performed is only binary. Therefore, more sophisticated models like convolutional neural networks are likely not necessary to achieve a sufficient accuracy. This project is to compare model accuracy based on training data, not to achieve maximum accuracy, so a classification accuracy between 60-90% is adequate.

References

[1] B. Rauenzahn, J. Chung, A. Kaufman, “Facing Bias in Facial Recognition Technology” *The Regulatory Review*, Mar, 20 2021. [Online]. Available: https://www.theregreview.org/2021/03/20/saturday-seminar-facing-bias-in-facial-recognition-technology/#:~:text=According%20to%20the%20researchers%2C%20facial,particularly%20vulnerable%20to%20algorithmic%20bias. [Accessed: Nov. 02 2022]

[2] A. Fawcett, “Understanding racial bias in machine learning algorithms” *DEV,* June, 8 2020. [Online]. Available: https://dev.to/educative/understanding-racial-bias-in-machine-learning-algorithms-4cij [Accessed. Nov, 02 2022]

[3] Darshan M. “How to use logistic regression for image classification?”. *Analytics Indian Magazine.* June, 5 2022. [Online]. Available: https://analyticsindiamag.com/how-to-use-logistic-regression-for-image-classification/ [Accessed: Nov, 02 2022]

[4] G. Galario, “Image Classification using Logistic Regression on the AMerican Sign Language MNIST”. *Medium*. June, 5 2020. [Online]. Available: https://medium.com/@gryangalario/image-classification-using-logistic-regression-on-the-american-sign-language-mnist-9c6522242ddf [Accessed: Nov, 02 2022]