Introduction

Machine learning algorithms that are taught with labelled data are used in many AI systems, such as facial recognition software. Algorithms taught with biased data have recently been shown to result in algorithmic prejudice. In the world of artificial intelligence, there is an array of shades towards certain biases. In case of this paper, it will relate to how AI can have a bias towards gender. This can cause a problem in many cases such as governmental issues, social and in the workplace.

Summary

Recent studies have shown machine learning algorithms to discriminate based on categories such as race and gender. In this case, work, designers provide a method for evaluating. In terms of phenotypic subgroups, there is bias in automated facial analysis algorithms and datasets.  using a system to characterize the gender and skin type distribution of two face analysis benchmarks, IJB-A and Adience, using the dermatologist-approved Fitzpatrick Skin Type categorization system. it was discovered that these datasets are primarily made up of lighter-skinned people (79.6% for IJB-A and 86.2 percent for Adience), thus we created a new facial analysis dataset that is gender and skin type balanced. Observing the dataset, one can see how the three commercial gender categorization methods analyse and find that darker-skinned females are the most misclassified (with error rates as high as 34.7 percent). For lighter-skinned boys, the maximum mistake rate is 0.8 percent. I

f commercial for fairness, significant differences in the accuracy of classifying darker females, lighter females, darker males, and lighter males in gender classification systems must be addressed immediately. (Buolamwini et al, pg.1, 2018) Bolukbasi et al. even shown that the widely used Word2Vec word embedding space reflects societal gender biases. The authors trained an analogy generator using Word2Vec, which fills in missing words in analogies. The man is likened to "X" who was a computer programmer, and "X" was a woman. completed with "homemaker," in keeping with the themed programming is related to a stereotype. Homemaking with women and men. The prejudices are hence likely to be spread by Word2Vec in any system that makes use of this embedding

Although numerous studies have looked into ways to make fairer algorithms and benchmarked discriminating in many circumstances (Kilbertus et al., 2017; Hardt et al., 2016b,a), just a few have looked into computer vision. Computer vision systems that perform poorly across populations, on the other hand, can have major consequences. Simple convolutional neural networks may be taught to detect melanoma from photos with accuracies comparable to specialists, according to Esteva et al (Esteva et al., 2017). However, without a dataset with labels for diverse skin features such as colour, thickness, and hair volume, it is impossible to assess the accuracy of automated skin cancer detection systems for people with various skin types. Similar to the well-known negative consequences of skewed clinical trials. (Popejoy and Fullerton, 2016; Melloni et al., 2010).

New contribution

There have been a few contributions to the study of Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on fairness, accountability, and transparency. These are, Phenotypic Labelling and improving the Intersectional Benchmark as well as Commercial Gender Classification Audit. Currently, evaluating gender classification performance necessitates breaking down the gender construct into several subcategories. To define genders " in this work, there was sex labels employed "male" and "female." Since the standards were examined, there has been a shift in gender classifications. These binary labels are used in categorization systems. An intersectional analysis also necessitates a with a dataset representing the defined genders with a dataset representing the defined genders with a dataset representing the defined traits that allow for subgroup accuracy analysis.

The other being Commercial Gender Classification Audit. Three commercial gender classifiers were tested. Male respondents were categorized more accurately than female subjects, confirming prior findings (Ngan et al., 2015), while lighter subjects were classified more accurately than darker people. All classifiers performed poorly on darker female individuals, according to an intersectional analysis. the positive predictive value of total classification accuracy, male classification accuracy, and female classification accuracy (PPV). We assess the true positive rate, false positive rate, and error rate (1-PPV) of the following groups: all subjects, male subjects, female subjects, lighter subjects, darker subjects, darker females, darker males, lighter females, and lighter males, in addition to the NIST Methodology.

This brings 3 questions. 1) How beneficial would it be, from a medical diagnosis assistance stand point would it be, to label genders along the spectrum and associated health risks that are involved 2) labelling genders along the spectrum could maybe make those who feel like they got left out/lost some validity. 3) For facial recognition even would you rather: no gender label, a binary (what someone *more* resembles, the extremes of the spectrum), or again labelling along the spectrum for the facial recognition machine learning to be able to associate with and pick up on . 1. think would be beneficial as it could create more comfort for a patient and then cater to health risk easier since there is more information . 2.There is a way around that someone could use a form which to submit and gets put into the database under "gender specific " (could probably use better title but point stands. It then could just be form based and more confidential.

3. The problem with the latter part would be then reinforce "fit this stereotypical characteristic to fit this gender" as machines like these can learn but can't critically think

Evaluation/critique

Through the methods that have been looked at in this paper there is a distinct number of problems with the works that were used . firstly, the method could’ve been more expansive in terms of what types of systems. While the types were appropriate somewhat it did not entirely account for all issues such as transgender skin changes after going on hormones and making the system inevitably not account for such issues and possibly misgendering a person. This socially can cause much stress to a transgender person and such a small thing from a system can have devastating consequences for their mental health. So, in turn there should have been other systems used that cover this especially considering the systems interchange sex and gender as if they are the same. The other problem is it does not take into account of other genders such as non-binary (while yes, the skin change will not always be different it is still important) and intersex for those who were not made to be one binary gender at birth. So, there is a necessity to do further research and look at it from these perspectives. This research could lead to improving the AI systems as there seems to be a lack of any systems that do not code in more then two genders. It could be possible to create a database for gender and upload them if needed such as in healthcare to take care of some of these problems.

There is a possibly a good impact on society if the before mentioned further research in this is done as this the sets a standard and expectation of what masculinity and femininity and androgyny should look like despite none of which should be required of anyone.

Conclusion

In conclusion there is possible benefits of this article but at same time it may damage certain people in different aspects and create a problem as the Fitzpatrick system isn’t the most conclusive baseline due to it usually checking for sunburn and while it did use the IJB-A skin test which is more inclusive it still seems problematic with what upcoming social issues could arise from the method and perspectives based on the evaluation. There is no denying the potential health benefits of identifying more genders along the spectrum for a benefit from a medical practitioner sense however does it just would it be enough to warrant benefit. Or would blood tests for hormone levels just yield much better medical data to act upon than gender identity.

References

Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on fairness, accountability, and transparency (pp. 77-91). PMLR.

Cathy O’Neil. Weapons of math destruction: How big data increases inequality and threatens democracy. Broadway Books, 2017.

Danielle Keats Citron and Frank A Pasquale. The scored society: due process for automated predictions. 2014.

Niki Kilbertus, Mateo Rojas-Carulla, Giambattista Parascandolo, Moritz Hardt, Dominik Janzing, and Bernhard Sch¨olkopf. Avoiding discrimination through causal reasoning. arXiv preprint arXiv:1706.02744, 2017.

Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639):115–118, 2017.

Alice B Popejoy and Stephanie M Fullerton. Genomics is failing on diversity. Nature, 538 (7624):161, 2016.

Chiara Melloni, Jeffrey S Berger, Tracy Y Wang, Funda Gunes, Amanda Stebbins, Karen S Pieper, Rowena J Dolor, Pamela S Douglas, Daniel B Mark, and L Kristin Newby. Representation of women in randomized clinical trials of cardiovascular disease prevention. Circulation: Cardiovascular Quality and Outcomes, 3(2):135–142, 2010