



The Myth of Machines Neutrality

Despite making all the claims, we all are not neutral, and we have biasness and discrimination among us. Then how can we say that things made by us will not have such traits and qualities? So, someone may ask, "Can programs also be biased?"

By **Zohaib Adnan** | Department of Computer Science | Published: September, 2026
✉ zohaibadnan31@gmail.com | Position: Software Engineer

Introduction

Humans are divided into many countries, cities, groups, colors, races, and breeds. But the problem is why we are divided into groups; some will say it is just for distinction, others will say it is purely to separate rich from poor or high-class people from low-class. There is nothing bad in above classifications, but the problem is the hate and discrimination that have evolved among these groups.

We used to believe that these traits are only limited to humans, and machines that don't have any heart, they cannot possess these qualities, but it is not true.

"I would argue that one of the major problems with our blind trust in algorithms is that we can propagate discriminatory patterns without acknowledging any kind of intent" — Cathy O'Neil

The bridge between Humans and Machines

Machines do not have a brain like humans; they don't have common sense and surrounding knowledge. If we see wet grass, we can conclude that there might have been rain at night, but computers and programs cannot simply conclude it. They only know what they are told and what they have been trained on. To train the programs, and especially AI, we provide it with our data, and they find patterns in it and make assumptions about it, and they build their knowledge from it.

The white Mask Incident

The biggest example of Above arguments is The White mask incident . Joy Buolamwini was a graduate student at MIT , while she was working on her project she discovered that the facial recognition systems (built by giants like IBM , Microsoft and Amazon) didn't work properly on her Dark skin and the systems immediately detected her face when she put the white mask on her face . Her main conclusions were

- All classifiers perform better on male faces than female faces (**8.1% – 20.6%** difference in error rate).
- All classifiers perform better on lighter faces than darker faces (**11.8% – 19.2%** difference in error rate) .
- All classifiers perform worst on darker female faces (**20.8% – 34.7%** error rate).



- Microsoft and IBM classifiers perform best on lighter male faces (error rates of **0.0% and 0.3%** respectively).
- Face++ classifiers perform best on darker male faces (**0.7%** error rate).
- The maximum difference in error rate between the best and worst classified groups is **34.4%**.

So we can easily conclude from above results that systems were performing better on male skins then female and also better on bright skins than dark skins , means If you are black female like Buolamwini you were going to get worsts results . It did not mean that systems were racist but the problem was in data , on which it was trained . Basically these were trained by and on white male with less number of females and dark persons so it did not work as expected and machine become racists in lighter terms.

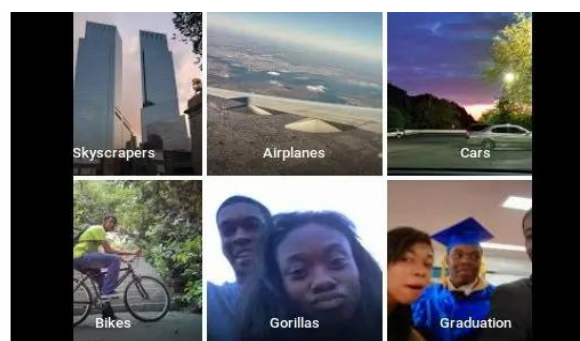
The above example best elaborate how can simple mistakes in training can make machines biased and prejudiced.

Value-Added Models (VAM)

Another example of machine based injustice were Value-Added Models which was used back in 2011 in School districts especially in Washington, D.C. , to evaluate teacher performance , an important event related to it was firing of a fifth-grade teacher named Sarah Wysocki . She was a very Good , loving and hardworking teacher , by any way of measurement she was best . But based on Value Added Models reports she was fired. The Algorithm was designed to evaluate teacher performance based on standardized test at the end of each year , The Algorithm use to decide a mean value , if students scores more then it teachers were awarded and vice versa. But here the developer made terrible mistakes and assumptions, according to Algorithm the student scores were only standards to measure the teacher performance completely ignoring the factors like compassion , sympathy , classroom engagement, and behavioral improvements, In simple terms the coded their own definition of success which was incorrect.

Later the found that teachers had intentionally given students more numbers to increase their performance report. we can conclude that the machine was not neutral in above case but as explained earlier, machines don't have emotions the only act according to their instructions. This is a Chilling example of how human mistakes can led to machine biasness.

The Google Photos incident (2015)



 **diri noir avec banan @jackyalcine** · Jun 29
Google Photos, y'all [redacted] My friend's not a gorilla.

TWITTER

Mr Alcine tweeted Google about the fact its app had misclassified his photo



Ever wondered why a giant like Google had to apologize in 2015? If you think it was just another apology for any other usual mistake, that's not the case. This story is about Google, when in 2015 Google launched its new product, the Google Photos App, which used an Artificial Intelligence-based algorithm that identifies what is present in users' pictures and, based on that, it used to make folders of it and categorize the pictures.

A 21-year-old Black programmer named Jacky Alciné was working on his computer, and suddenly he saw that his and his Black girlfriend's pictures had been labeled as "Gorillas." He took a screenshot and posted it on Twitter, asking questions to Google about what kind of image data they had collected that resulted in such a racist and horrific categorization.

Then within a few hours, Google apologized for it and told him they would fix it immediately; although they removed some keywords, the issue was never fully resolved.

"We're appalled and genuinely sorry that this happened. We are taking immediate action to prevent this type of result from appearing. There is still clearly a lot of work to do with automatic image labeling." — Google spokeswoman in 2015

Coded Bias

This is a documentary that was released in 2020, and in it, the filmmaker explained many events related to machine biasness. Some of the main ones are as follows:

- Value-Added Models (VAM).
- The white Mask Incident.
- The UK Police and Live Facial Recognition.
- Amazon's Scrapped Hiring AI.

Conclusions

We used to think that AI will become too smart one day and will act against humans, but reality is much more different from it.

"The danger is not that AI will become unlike us, but what is horrific is that it will become exactly like us"

Machines learn from our data and our history, and they eventually execute our biasness with horrific and terrifying accuracy. The first step toward betterment is to understand our programs are inherently flawed because their creators are flawed. The code is the same as the humans who write it.



About the Author

My name is Zohaib Adnan , I am software engineering student and an author , for more details and connectivity you can contact me any time at zohaibadnan31@gmail.com

Keywords: Algorithmic bias; Software Engineering Ethics; IoT; Pakistan Tech

References

- [1] J. Buolamwini and T. Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," *Proceedings of Machine Learning Research*, vol. 81, pp. 77-91, 2018. [Online]. Available: <https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>
- [2] B. Turque, "Creative, motivating — and fired," *The Washington Post*, Feb. 04, 2012. [Online]. Available: https://www.washingtonpost.com/local/education/creative--motivating-and-fired/2012/02/04/g!QAwzZpvR_story.html
- [3] J. Kasperkevic, "Google says sorry for racist auto-tag in photo app," *The Guardian*, Jul. 01, 2015. [Online]. Available: <https://www.theguardian.com/technology/2015/jul/01/google-sorry-racist-auto-tag-photo-app>
- [4] BBC News, "Google apologises for Photos app racist blunder," *BBC News*, Jul. 01, 2015. [Online]. Available: <https://www.bbc.com/news/technology-33347866>