

Bias in AI Essay

xchs48
xchs48@durham.ac.uk

I. MY OPINIONS ON A REVIEW INTO BIAS IN ALGORITHMIC DECISION MAKING

The report begins by making it clear that bias in algorithmic decision making is a problem that is yet to be properly addressed. However it makes the case that the best time to act on changing this is now, before bias becomes entrenched. This is important as this report is a government report, intended for those in a position to implement reforms and legislation.

The recommendation to act now will help persuade politicians that changes have to be made quickly and should be a priority for lawmakers. Legislation often lags behind technology in such developing fields as they might be seen as unimportant or are not fully understood by those in power. Therefore the emphasis on immediate action helps bring forward when legislation to help may be put in place.

Also of note is how each key area that algorithmic decision making is being used in is given an overview section and a number of recommendations detailing how things could be improved in that sector. This use of a simpler and easier to read section helps those outside of the field understand the issues, and clearly explains what steps have to be taken to mitigate these issues helping to provide a framework for new laws that could be implemented. As most people do not have a strong understanding in algorithmic decision making this clear and well explained structure is important for helping the government and the public understand these issues and how they can be solved.

It is highlighted how it is important for algorithmic decision making to gain the public's trust. This is because people will find it hard to accept results from algorithmic decision making if they do not believe they are unbiased. On this point it brings up the example of how public opinion changed following the exam results last year, showing how this applies in practice, and how much of an impact results people feel are biased can have. This means that for the field to develop and be accepted for mainstream use it has to appear trustworthy to those affected by it.

In conclusion the report provides important information for lawmakers on how algorithmic decision making should be treated by the law concerning bias, and stresses the need for action in making this field fair and unbiased. This will hopefully help regulate algorithmic decision making, leading to results not burdened by bias and that can be trusted by the general public.

II. THE EVOLUTION OF DISCRIMINATION-AWARE ARTIFICIAL INTELLIGENCE

The field of artificial intelligence is becoming more mainstream as it becomes easier to implement and more useful. As this happens it becomes more important to identify and reduce bias in these systems so they are fair for everyone.

As the field of Artificial Intelligence in its current form is still relatively new there is a lack of legislation to force companies to combat it [1]. Therefore it is still up to private businesses whether they want to spend the time and money to ensure their AI models are bias free. The lack of standards means that companies' implementations of bias reducing algorithms will lack conformity and an ability to assess how effective they are. Adoption of industry wide standards or government legislation are therefore a deciding factor in the effectiveness of reducing bias in AI, as otherwise it does not matter how effective the techniques are if they are not implemented in the systems where they can make a difference.

A lot of good algorithms and tools for reducing bias have been created, allowing easier and more effective implementation of techniques. These can be used to successfully combat bias, helping fairer systems be created. This shows that research is being put into this field, and advancements are being made quickly. If this continues it will only become easier to reduce bias in your model, hopefully leading to these techniques being more readily considered for including in models.

However, while a lot of research has been done into combating some types of bias in models, others areas are underdeveloped. More research has to be done in these areas that are behind as it is important for bias to be removed for all types of AI implementation [2]. Research has to be better spread for this to happen, with less of a focus on classification algorithms. I believe this research will happen with time, but it will take a while for all fields of AI to be fully caught up. Until then bias mitigation techniques may be fairly weak in some areas. It will therefore be important to note how effective the techniques for certain algorithms are and to take this into account when designing an AI system.

While progress has been made, the field of combating bias for AI is behind where it needs to be for how important AI is in the modern world. However I believe that the field can and will be developed further, and be an important part of future AI projects, making the world a fairer place.

REFERENCES

- [1] "Review into bias in algorithmic decision-making", November 2020.
- [2] N Mehrabi, F Morstatter, N Saxena, K Lerman, and A Galstyan, "A Survey on Bias and Fairness in Machine Learning", September 2019.