

PLAGIARISM SCAN REPORT

Words 452 Date April 22,2021

Characters 2934 Excluded URL

8%

Plagiarism

92%

Unique

2

Plagiarized
Sentences

24

Unique Sentences

Content Checked For Plagiarism

With the surge of the fourth wave of the Industrial Revolution, machine intelligence, blockchain-based decentralized governance, genome editing and AI in healthcare are among the top trending arenas, and have grabbed the interest of researchers and enthusiastic techies. But despite their enormous impact on industries, they also pose new ethical challenges. Curious?

Biases in the psychological world are quite common. Humans are more or less always inclined towards a certain object or an opinion. But can one agree with the thought of an algorithmic model favouring something over the other? Being able to generalize a problem over a set of inputs represents the key characteristic of a machine learning algorithm, meaning it must be able to correctly predict the outcome for new data, based on insights gained from previous data. But if the incoming data contains unseen features, the algorithm will have trouble identifying what this new data is. If this all sounds a bit abstract, here is a quick example.

Imagine developing an image classification model, for identifying a cat or a dog. In the training phase, it is fed hundreds of thousands of labelled images of dogs of different breeds, but very few of the different breeds of cats. Thus, it is very likely for the model to misclassify an image of a persian cat as that of a dog. The model is not generalizable enough to all cat breeds, because it has not been trained with the sufficient images of cat breeds. This represents a bias in an image classification problem.

The root cause

Data imbalance is one of the major factors in introducing bias. In 2016, Microsoft unveiled an AI-based conversational chatbot on Twitter to interact with people. However, within a few hours of its release, the replies became quite offensive, and loaded with racist messages. The chatbot was trained on anonymous public data and had a built-in learning feature, which led to a coordinated attack by a group of people, allowing introduction of racism in the system. The episode was an eye-opener for many, of the potential negative implications of unfair algorithmic bias. Class imbalance is the another leading issue in many classification problems. Researchers and developers may also share some blame, termed human bias.

Prevention Measures

Awareness and good governance can help prevent machine learning bias, by cultivating the best practices to mitigate it. Selecting a representative sample will counteract common types of machine learning and artificial intelligence biases. Also, continually monitoring systems as they learn and execute can help ensure that biases don't sneak in over time. Additional resources, such as Google's What-if Tool or IBM's AI Fairness 360 Open Source Toolkit, to examine and inspect models, will be a step towards trusted AI.

Sources

Similarity

9%

<p>Bias in AI systems and efforts to fix it by Rugare Maruzani ...</p> <p>If this all sounds a bit abstract, here is a quick example. Imagine you want to train an ML model that classifies if a photograph has a horse or a dog in it. ... The model is not generalisable enough to all dog breeds because it has not been trained with images of other dog breeds.</p> <p>https://towardsdatascience.com/bias-in-ai-systems-and-efforts-to-fix-it-318798f5b39f</p>	
<p>The Role Of Bias In Artificial Intelligence</p> <p>· The chatbot was trained on anonymous public data and had a built-in internal learning feature, which led to a coordinated attack by a group of people to introduce racist bias in the system. Some...</p> <p>https://www.forbes.com/sites/forbestechcouncil/2021/02/04/the-role-of-bias-in-artificial-intelligence/</p>	5%