



## Algorithmic Biases in Artificial Intelligence is for Real

As companies increasingly adopt Artificial Intelligence and Machine Learning in our public and private sphere, complex algorithms are set to work on large and updated data sets. The speed of processing this data resulting in actionable insight is the advantage that's driving this billion-dollar industry. But the inherent risks in this fast-paced technology is the Algorithmic bias that can trigger costly errors in the long run. If unchecked this problem can escalate and pull projects off rail.

By 2021, IDC research firm predicts that enterprises will allocate around \$52.2 billion annually on AI related products. This will be a tremendous savings in their CAPEX & OPEX expenditure and substantial gains from that investment. Most of that will happen from reducing headcount but far more will come from enormous efficiencies in matching solutions to the problem and product to the customer.

With this much at stake, it's imperative that organizations tread carefully when handing over their future to this new technology. It calls for more protective measures from the potential drastic consequences of the AI biases.

### Human Biases and AI

Algorithms and artificial intelligence (AI) are intended to minimize human involvement and emotions in data processing that can be compromised by human error and would expect the data to be completely sanitized. However, any human bias or error in collecting the data going into the algorithm will in reality be exaggerated in the AI output.

The AI algorithmic decisions are only as good as the data that humans feed them. As their builders are learning, the data used to train deep-learning systems isn't neutral. It can easily reflect the biases whether consciously and unconsciously of the people who design it.

A typical use case where human biases could cloud the actual output of the AI algorithm is where the insurers based on the algorithm, charge extra premium to ethnic, minority race irrespective of location and other factors.

Identifying and mitigating bias in AI systems is essential to building trust between humans and machines that learn.



As AI systems find, understand, and point out human inconsistencies in decision making, they also help us identify ways in which we are biased or partial and lead us to be more practical and adopt an unbiased, egalitarian view. In the process of recognizing our bias and teaching machines about our common values, we may improve more than AI.

### Some of the potential biases that perpetuate in machine learning are:

- **Interaction bias:** If we are teaching a system to learn to recognize what an object looks like, then what we teach it to recognize is skewed by our interpretation of that object and the biases that we have as a user. The algorithm will only learn and build upon that basis and will reflect the opinions of the people who train them.

- **Latent bias:** If you're training your program to recognize a doctor and your data sample is of previous famous physicists, the program will be highly skewed towards males.
- **Confirmation bias:** Consider the classic example of Facebook users and their data feed. The algorithm only displays posts based on their shares and likes. It analyses the data feed to present new content that matches what it confirms to the user's set of existing ideas.
- **Similarity bias:** This is simply the result of an algorithm that's designed to output the result based on "similar to" either the requested queries or existing documents. Example is that of a news feed that displays similar stories that tend to confirm and corroborate each other.
- **Selection bias:** The data used to train the algorithm over represents one population, making it operate better for them at the expense of others.
- **Conflicting goals bias:** The system is designed for very specific tasks, but the result will be influenced by the user's bias reinforced by their ideas of stereotypes. Imagine the system that is used for displaying the most accessed job post in the website resulting in displaying the job post that has the highest number of clicks. Say, for a secretary job post, this will be most accessed by women rather than men. This is due to the typical individual's knowledge of a stereotype associated with the secretary job that leads to greater clicks.

- **Data Driven bias:** The output of the AI systems is driven by the millions of data that's fed to the systems. If the training data itself is skewed, then the result will also be equally so. Though this can be unintentional, but they demonstrate the possible negative repercussions that occur from learning from a skewed data.

## Can biases be addressed in machine-learning algorithms?

While its nearly impossible to eliminate algorithmic biases in AI and machine learning, organizations can implement measures that mitigate the algorithmic biases in decision making process.

The Zero Incident Framework™ is built keeping the below factors in mind when designing the AI algorithms. This enables us to provide an inherently unbiased platform for organizations to deploy and perform smarter IT Operations.

- Machine learning algorithms are built by humans and they need to understand an algorithm's shortcomings and design them such that the output is not invalidated by algorithmic bias. They must understand the operating procedures, safety practices and best rules.

- Available data samples for teaching the algorithms must be unbiased by historical data. This is a vital step that will ensure the data generated and collected through a controlled environment is fresh and unbiased.
- Not every scenario is feasible for using machine learning algorithms. Organizations should know when to use and when not to use them to truly leverage its benefits. While algorithms offer speed, accuracy and convenience, the regular manually crafted models, such as decision trees or logistic regression or even human decision-making approaches provide more flexibility and transparency.

The future of AI based learning is in unbiased algorithms. It's still work-in-progress with great potential to benefit humans. What's important is to recognize that those biases exist, and that policymakers try to mitigate them.

GAVS Technologies expertise in AI based solutions is driving organizations in their endeavor to maximize their operations and business prospects through AI initiatives.

## About GAVS

GAVS Technologies is an automation-led digital transformation company with focus on AI, Predictive Analytics and robotics-led Infrastructure Management Services. GAVS' IP led solution, Zero Incident Framework™ (ZIF) is an AIOps solution that provides a 360-degree view of enterprise IT health, proactively detects incidents before they occur and remediates with minimal human interference. By focusing on eradication and proactive remediation of the incidents, ZIF enables organizations to trend towards a Zero Incident Enterprise™. GAVS is committed to improving user experience by 10X and reducing resource utilization by 40%.