Combatting Bias in Machine Learning Algorithms?

For Recruiting

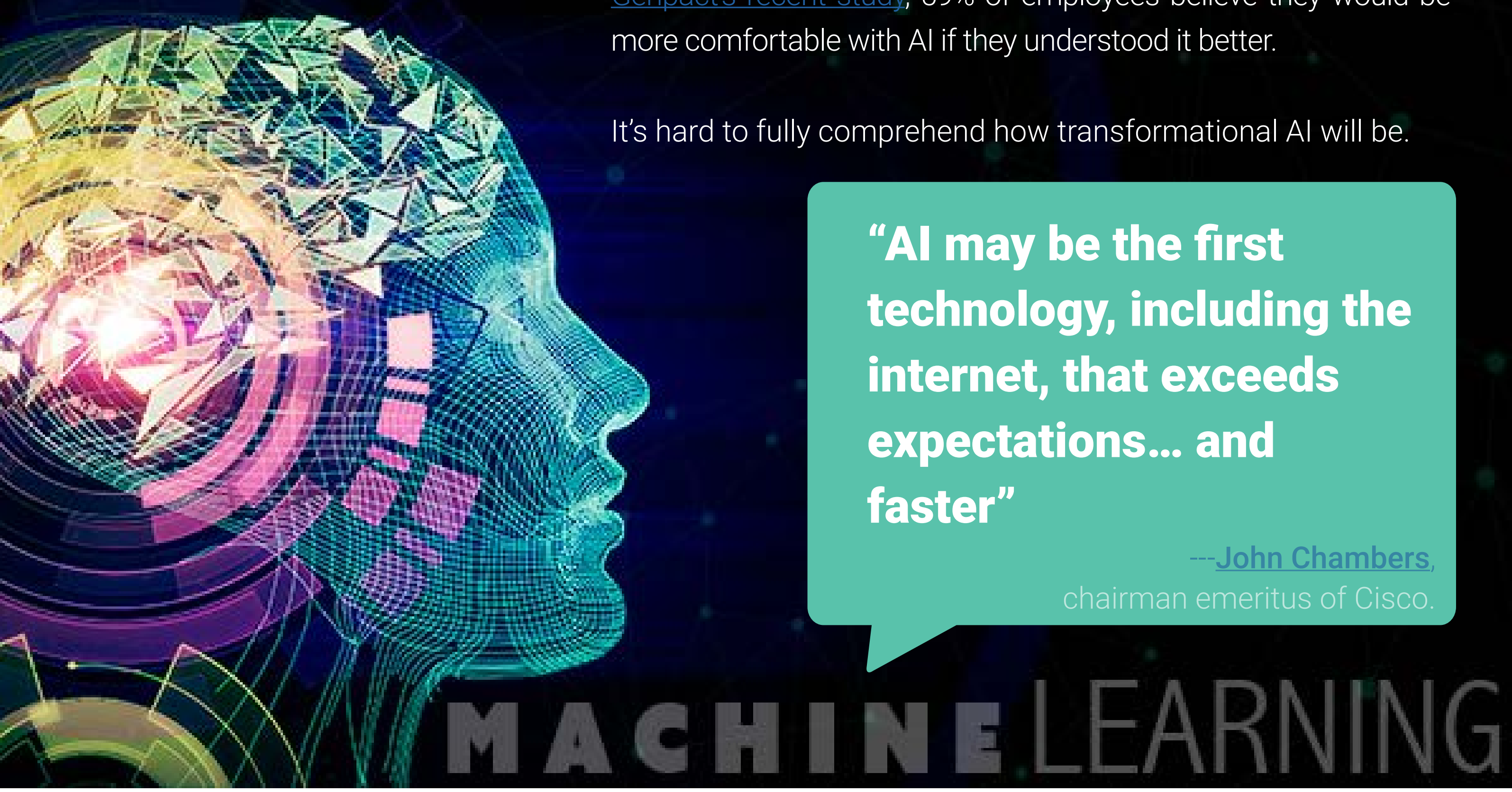
Machine learning in its rawest form

AI and machine learning have permeated nearly every area of our lives over the last half decade. One of the most controversial questions throughout the technology's evolution has been, "How do we prevent machines from learning the worst of our own traits?" This question has been strongly debated, heavily researched, and continuously tested for years and years. Is it even possible to build AI that doesn't reflect the underlying prejudices that are built into our social systems? Are technology vendors lying when they say their machine learning software eliminates bias?

For most ordinary people, machine learning lives in a black-box. Most feel like they can't understand its inner workings, so it's difficult to trust and depend on for critical decision-making. According to Genpact's recent study, 59% of employees believe they would be more comfortable with AI if they understood it better.

It's hard to fully comprehend how transformational AI will be.

> **"AI may be the first technology, including the internet, that exceeds expectations… and faster"**
>
> ---John Chambers, chairman emeritus of Cisco.



sales@goarya.com 1-800-256-2925 www.goarya.com

But the rapidly penetrating bias of machine learning, and the lack of transparency and diversity in its programmed algorithms has made it difficult for business leaders to accept and adopt with confidence.

Most know by now that Amazon scrapped their AI recruiting tool because of biases toward women. How did a giant company like Amazon let biased software slip into their recruiting strategies? Can we trust AI systems to make unbiased decisions?

Define the bias in your recruiting process

Humans, in general, are biased by nature. Therefore, the outcomes of the processes we build are inherently biased. Your past recruiting processes? Biased. Your present recruiting processes? Probably biased. You may not even know it. Our bone-deep, intrinsic prejudices combined with our past experiences reveal themselves in our decision-making without us consciously realizing it.

Because it's so difficult for us to recognize and understand our own conscious and unconscious biases, it's even more difficult not to feed them into technologies. When that happens they are then deeply embedded, relearned, and reinforced in a company's decision-making.

There are two distinct types of bias: systematic and statistical. Systematic bias can be broadly defined as "AI and machine learning models feeding datasets that produce erroneous or inefficient results due to inherent biases or insufficient data diversity." Statistical bias produces erroneous or inefficient results due to incorrect calculations.

Here is an example:

Humans have repeatedly chosen to hire white males at a company or in a department when there are a plethora of competing candidates of varying backgrounds and demographics with equal or greater aptitude.

When this historical data is fed to an AI recruiting tool, it predicts that white males are the best candidates. There is zero statistical bias influencing those predictions because the machine is using the data correctly. The bias is systematic. It is a result of the machine accurately processing inherently biased data.

In this example, the AI has interpreted the data about hiring success correctly but has generated biased results because of the information that humans have provided it.

So, the question is, **how do we feed quality data into our technologies to minimize systemic bias within the results?**

Develop ethical policies and objectives

According to research by IBM, more than 180 human biases have been defined and classified, any one of which can affect how we make decisions.

In order to keep these biases from creeping into the foundation of our technologies, organizations have to set objectives and/or set precision rules before implementation of a machine learning model to challenge fair and unbiased machine learning outcomes.

Let's say you set an organizational objective to hire a more diverse workforce. Ask yourself, is your objective to create a diverse workforce for compliance purposes or a diverse workforce for better outcomes? It's important to define that before you measure success, and it's your job as an organization to set objectives and define the biases you don't want linked to your business and recruiting strategy.

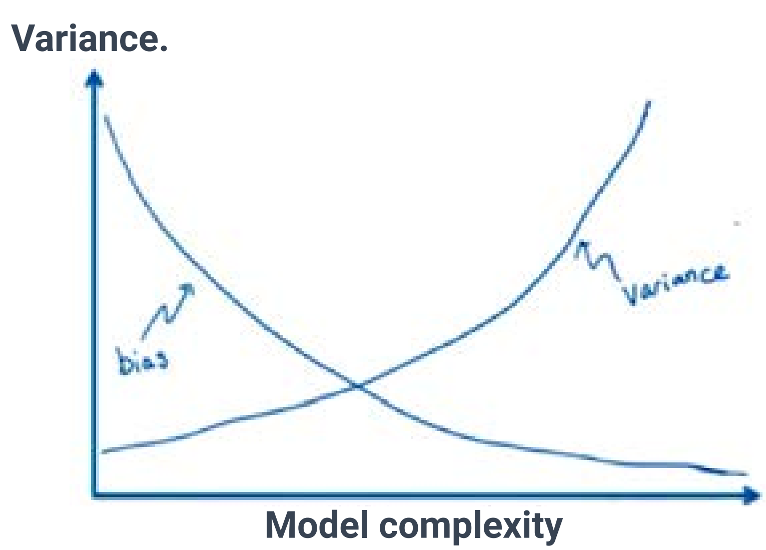
It is a mistake to immediately believe our technologies are biased without understanding why. Are you simply basing that on average society biases? Is the technology bringing back more white candidates than black? More men than women? More candidates from a certain university? Is there a reason for this based on the data the system was measuring?

How to handle Biases?

Here are just a few important points to think about before settling on a solution:

1. Do you have enough adequate data (including data diversity) to train and test your algorithms? If not, augment more generalized data within your industry and gather competitive data.
2. Gain deeper understanding of the features the algorithms are working with and eliminate the obvious ones that could contribute to biases (like gender information).
3. Identify the right algorithms that are in-line with your organizational objectives. Some algorithms have better capabilities to deal with biases.

Some models have a relationship between complexity and variance:



LASSO Regression:

- LASSO stands for Least Absolute Shrinkage Selection Operator. Linear regression uses the Ordinary Least Squares (OLS) method to estimate the coefficients of the features. Interestingly, LASSO incorporates statistical bias to improve model accuracy. That is, as part of the regularization technique (or the introduction of bias), the prediction accuracy is found to be improved by shrinking the coefficients of one or more of the insignificant parameters/features to bare minimum or near-to-zero (Ridge regression) or zero (LASSO regression). Ridge regression helps in estimating important features. However, LASSO regression helps in confirming the most important features as the non-significant features' coefficient is set to 0.
- An additional value of LASSO is its transparency. The relationship between the variables is clearly stated, and the coefficients modeled can be seen after training. In other words, you know exactly how much each variable included is effecting the outcome or prediction. This is opposed to more complex models such as random forests or neural networks which can learn feature interactions for you.

Random Forest:

Random Forest, an ensemble modeling technique, is used to determine feature importance using the following techniques:

- The depth of the attributes in the decision tree can be used to determine the importance of the attributes/features. Features higher in the tree could be thought of as affecting a significant portion of the total samples used to learn the tree model. Thus they are termed features of higher significance.
- Permuting the value of attributes/features and determining the performance averaged across all the trees. If the attributes/features are highly significant, the change in accuracy would be high.
- Random Forests are an example of a model that can produce higher quality predictions, if you trust the quality and diversity of your data. Random Forests, like neural networks, can learn feature interactions. And while this can be incredibly beneficial, it does possess its threats. Consider the case where your model includes both university and major. Alone, both of these features in LASSO can be fairly reasonable, however when combined it is possible that your model will instead be learning that, because you went to this university where you

studied this major, it is likely you are female and females tend to rarely enter this profession or this job. The reality is, other features may indicate that this candidate is perfectly suitable for the role, but if gender had been included as a variable it could potentially overshadow features we would say are more important (certifications, experience, etc.). The takeaway is that complex models can potentially learn features you intentionally did not provide.

4. Set the precedent of appropriate precision rules on top of models to achieve desired results

   Let's say you have one hundred candidates, and your organization needs 20% of these candidates to be diverse in order to meet a compliance or results-driven objective. You'd set a precision rule on top of the final outcome to make that happen. This could be done by wrapping an optimization criteria around your organization's machine learning systems where you incorporate diversity constraints.

5. What is the deeper philosophy behind the models and AI? This could profoundly influence your selection of algorithms.

For example, you can look at a car as an automobile with some type of engine, or you can look at a car as a way to transport people. These are different philosophies. Underlying philosophies drive data scientists to sanitize the data and produce models that get the best results. It is up to the value provider to describe, "what is your philosophy for producing this outcome?"

Companies like IBM are beginning to develop open-source models to identify cognitive biases in machine learning algorithms. Their mission reads, "By building open source models and training data with increased transparency, providing more opportunity for hidden biases to be uncovered." This may not solve the problem of cognitive bias in machine learning as a whole, but it opens the doors toward collaboration and innovation in this space.

So it is the combination of philosophy, objectives, algorithms, and a healthy data ecosystem that can handle the biases for desired outcomes.

From the words of recruiting leader, Katrina Kibben, "AI is not a panacea. It's a strategic complement to the human aspects of the process, including the one element that will likely always rest with people – deciding, or not, to make the hire."

As more companies continue to bet big on AI to help them overcome unconscious biases, it is crucial that HR leaders work with AI throughout the hiring process, rather than relying on it to solve institutional bias. AI is generating unlimited opportunities for the use of data and new analytic techniques in the workplace, which is why they must be ethically leveraged and implemented", states Infor's vice president of science applications, Jill Strange.

These tools weren't made to make decisions completely on their own. Once organizations incorporate their objectives and results-driven goals, they not only create better decision-making criteria for their team, but they develop better recruiters who actually trust in their data-driven results.

AI and Machine Learning are rising to the top of every company's tech wish list

As companies begin to trust AI and machine learning vendors, there's no stopping the growth of the industry. Companies of every size hope to use machine and deep learning-powered technology to mine their data and find the hidden gems that are too deep for human eyes to identify.

IDC predicts worldwide spending on cognitive and Artificial Intelligence systems will reach $77.6B in 2022. Gartner predicts the business value created by AI will reach $3.9T in 2022. **More than three times the $24.0B forecast for 2018.**

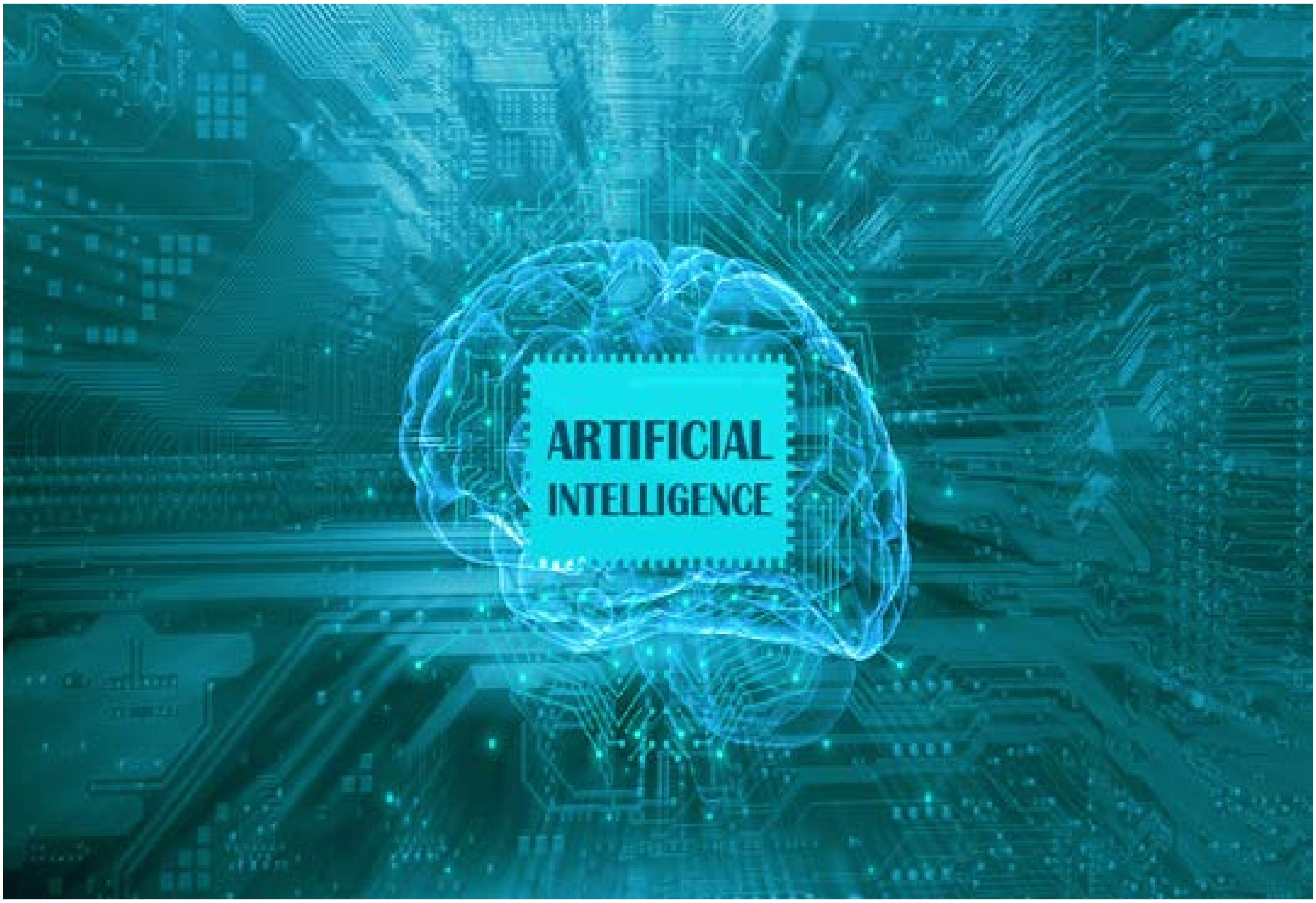
The team at People Matters adds, "Organizations still need to understand their talent needs better to fruitfully utilize the potential of AI. It is only when they have clear, solid measures of human performance that they can build meaningful models to predict performance and quantify a person's suitability for a role or job."

"AI used properly in the HR hiring process can get past the clone-seeking of yesteryear's recruiting applications that produced a plethora of corporate environments lacking gender, racial, and creative diversity. But AI needs data, and it needs to access more information than who was hired in the past and how to get more of him or her."

---Katherine Jones,
High-tech market analyst.

Don’t let the fear of bias stall your AI efforts

Don’t be panicked about bias. Exercise caution with objectives and compliance, and make sure that you’re not stepping on legal statutes. Make sure you have compliance reports ready to be examined or put into place, so that you don’t have to worry. But don’t let the fear of biases prevent you from innovating and investing in AI and machine learning initiatives. You need to strike a strategic balance between compliance and innovation.



Human hands are touching data, so it’s impossible to not have a trace of bias in its results, but consistently monitoring and dedicating resources to reviewing the intelligence is how to ensure AI and humans are working simultaneously instead of in silos.

About Arya by Leoforce:

We've recognized that it's SO important to be transparent about what our AI is actually doing behind the scenes. As an AI company, it's critical to assess our algorithms on a consistent basis for bias, clarity, and data security and begin to refine as needed. **Arya monitors, refines, and tests each aggregated data point for fairness, accuracy, and relevance on a consistent basis to provide the freshest and most reliable candidate, job, and industry data for your recruiting team in real-time.** Providing trust to our clients is our number one goal as a technology company. And that starts with transparency around our data.

ARYA™

Want to learn more about our AI Recruiting Platform?

Get your free demo here

Or call: 1-800-256-2925