Building an AI framework for fair hiring: A U.S. employer puts antibias first

IBM teams with a large corporation to operationalize trustworthy AI at scale

Data science can quickly turn data into insights and those insights can lead to decisions. And sometimes, the results are unwittingly spoiled by bias and drift, causing mistrust. This problem undoubtedly hampers AI adoption and can negatively impact people's lives and a company's reputation.

Take hiring decisions.

Tools or recruiting systems that screen candidates have long demanded attention; as research has demonstrated, they can reflect historical discrimination based on the datasets.

Sensitive features such as gender, ethnicity, and age, even if not included in AI, could have influenced the training of the data, the source of the data, and even how the data got to the AI from a training dataset. In other words, even if there's no intent or access to those features in the beginning, those perceptions can lead to incorrect decisions.

A growing concern about AI's trustworthiness has provoked worldwide conversation among data leaders and business leaders alike about how to improve the practices of trustworthy AI and govern it across the AI lifecycle.

How do we understand what AI models are doing?

How do we ensure AI accuracy and fairness?

How do we speed up production and adoption of AI models?

Can we trust the output?

According to IBM, if a business is involved in making decisions on automation that's driven by AI, it needs to be transparent. The business must know it's making decisions that align with company policy — and that people who are making the decisions based on AI can trust it.

One major U.S. company was eager to tackle the problem on a large scale and turned to IBM for help. Within this corporation's mandate to focus on social responsibility has been an effort to drive more workforce diversity and inclusion. When it came to its hiring practices, it was critical that this employer ensure fairness and trust was in place within its AI and ML models – especially when it came to attracting and recruiting talent.

With over 1,000 data scientists in its ranks, this industry leader has traveled far on its AI journey. Hundreds of ML models were in production, but what it lacked was an enterprise solution that assured that models could be trusted in a socially responsible manner.

Data science leaders wanted to be able to translate the models' decisions and results easily — in a way any hiring manager could understand. It wanted to establish fairness by accelerating the identification of any bias in hiring and "explain" decisions made by AI models. The company also knew it needed to operationalize AI governance to get more of its business users on board – so it set out to find a solution that could achieve all of these things.

The answer was IBM Watson® Studio(TM), a AI monitoring and management tool within IBM Cloud Pak® for Data that filled a much needed gap. Once IBM's Data Science and AI Elite team showed how the product could consistently manage AI models for accuracy and fairness, IBM's Expert Lab services came in to drive the ongoing teamwork needed to reach the corporation's goals.

Over 90% of organizations say their ability to explain how their AI made a decision is critical.

So what's the next step to put trustworthy AI into practice?

Since partnering closely with IBM, the company has been tapping IBM's Expert Lab services to implement IBM Watson Studio on Cloud Pak for Data in several use cases, relying on IBM's expertise for this area of the AI lifecycle. The partnership has resulted in the creation of a enterprise framework that can operate within the scale of the enormous organization. Today the customer has all the capabilities it needs to manage aspects of bias, fairness, accuracy, drift, explainability and transparency in its use of AI and machine learning.

Now, the company is proactively monitoring for and mitigating bias in its hiring processes. Because automation has reduced the workload within DevOps, the company's data scientists can focus more on the new model development and refinements.

Today, companies across all industries have a clear opportunity to harness data and AI to build effective and scalable solutions while eradicating systemic racism and structural inequality. And there's no denying the fact that there's a relationship between higher growth and the ability to scale AI with repeatable, trustworthy processes. According to a January 2020 Forrester Consulting study commissioned by IBM, *Overcome Obstacles to get to AI at scale*, the companies who are the fastest growing in their industries are over 6x times more likely to have scaled AI.

There's no better time to address the societal relevance of AI and the need for a trustworthy AI framework based on ethics, has governed data and AI technology, and is rooted in a diverse and open ecosystem.

Accelerate your AI journey with a prescriptive approach.