

“Ethical Considerations in ML” Module Transcript

Chapter 1

Intro, Topics Covered, & Learning Outcomes

Hi all. My name is Hayley Bresina, and I'm a client enablement specialist on the One AI team here at One Model. In the "What is One AI?" module, we touched on that One AI is a fully explainable machine learning platform that was designed specifically to prioritize transparency and ethics.

Its purpose is to enable organizations to make well informed and ethical decisions about their people by ensuring transparency at every step. In this module, we will explore the ethical considerations surrounding machine learning and how One AI addresses these concerns through its design and application.

Specifically, we will cover an overview of the ethical considerations in machine learning, how One AI as an ethics first platform addresses ethical concerns, and why ethical machine learning matters and is beneficial for business.

After completing this module, you will gain a comprehensive understanding of why ethical considerations are critical in the development and application of machine learning models, especially in people analytics. You will understand how One AI's features, such as transparent reporting and bias mitigation tools, support ethical use of machine learning within your organization. You will gain practical knowledge on how to implement best practices to enhance privacy, ensure fairness, maintain data integrity, and improve transparency and explainability and machine learning models. And you will learn how to articulate the business and organizational benefits of implementing ethical machine learning.

Chapter 2

Overview of Ethical Considerations in ML

Section 2 - Overview of Ethical Considerations in Machine Learning

Due to the nature of our work in people analytics, the use of machine learning, which involves analyzing employee data to gain insights and inform workforce management decisions, raises several ethical concerns that demand careful consideration and intentional planning. We will discuss some of these major ethical considerations now.

First and foremost, considerations around privacy and confidentiality are crucial. Machine learning models typically require large datasets, which in people analytics often contain very sensitive personal data. There should be robust measures put in place to protect data against unauthorized access and ensure that the data is only used for intended purposes.

Next, we must consider bias and fairness. Machine learning algorithms can perpetuate or even exacerbate biases present in the training data, potentially leading to unfair outcomes for certain groups based on gender, race, age, or other characteristics.

It's important to actively identify and correct biases in datasets to ensure fairness in decisions like hiring, promotions, and terminations.

Transparency and explainability are essential. Depending on their use cases and problem domains, decisions made by machine learning models can significantly impact employees' careers and well-being, as well as the organization's strategy overall.

Therefore, these models should be transparent and their decisions explainable.

This involves explainable data predictions and algorithms to aid in understanding how and why decision was made by the model, which can be challenging with more complex algorithms and black box models.

We also need to consider data accuracy and quality.

Model predictions are only as good as the data they're trained on. Inaccurate or poor quality data leads to incorrect conclusions, which can be harmful. Ensuring data integrity and continuously verifying data quality is essential.

And finally, regulatory compliance is key. Organizations must comply with relevant laws and regulations related to data protection, employment, and anti discrimination, such as GDPR in Europe. Staying ahead of evolving technology and thinking proactively about future regulations can prevent disruptions in your machine learning practices as laws catch up with technological advancements.

Considering these ethical concerns is not only about mitigating risks, but also about building trust with employees and creating a fair and inclusive workplace.

This involves continuous dialogue with stakeholders, ethical review processes, and adapting policies as needed.

Considering the ethical concerns is only the first step. In the next section, we will discuss how One AI and your organization can take action to address these concerns.

Chapter 3

One AI Addresses Ethical Concerns

Section 3 - How One AI, as an Ethics-first Platform, Addresses Ethical Concerns

Having outlined the key ethical considerations in machine learning, we will now explore how One AI was specifically designed to address each of these concerns.

I will also provide guidance on best practices that organizations can adopt to ensure their machine learning initiatives are both effective and ethically sound.

First, let's revisit privacy and confidentiality. One AI exclusively uses data that has been loaded into One Model and validated by the team, ensuring that data remains within organizational boundaries and is not mixed with external datasets.

Data does not leave One Model unless it's downloaded for external use by an authorized user.

Additionally, access to the One AI machine learning tool is safeguarded by an application access role, which is granted exclusively by site administrators.

Best practices include conducting regular audits to monitor which users and roles have access to the One AI tool using only data that is necessary and relevant to the specific problem domain of the model and routinely auditing data access and model training processes.

Returning to bias and fairness, One AI includes tools to select protected classes, debias the model by removing disparate impact, and utilize fairness metrics. This helps in actively identifying and mitigating biases in the training data. Best practices include routinely reviewing and updating the training data to ensure it accurately reflects a diverse workforce.

Additionally, recognizing that if the training data contains biases, the model will inevitably learn and potentially amplify them, leading to biased predictions and decisions is crucial.

Therefore, ensuring the training data is unbiased is necessary for fair outcomes.

Calling back to transparency and explainability, One AI provides detailed transparent reports each time the model is run. The exploratory data analysis report illustrates how variable selection was determined and allows users to conduct variable and correlation analysis to better understand the model dataset.

Paired with the results summary report, which details the model's performance and configuration, as well as feature selection, explainability of the model's predictions is enhanced greatly. Best practices include avoiding black box machine learning solutions, regularly engaging stakeholders, only deploying models that make sense and are easily explainable, and carefully analyzing model reports and performance. Additionally, visualizing model outcomes in One Model storyboards can help make abstract concepts more concrete and understandable.

Reflecting on data accuracy and quality, One AI restricts model training to an organization's validated historical data, which enhances the accuracy and relevance.

As your organization loads more data sources into One Model or corrects errors with an existing data, This new data can easily be added to the model dataset. One AI also employs rigorous data pre processing where existing data is cleaned and transformed automatically excluding data with issues such as being overly null or prone to data leakage.

Best practices include rigorous data cleaning processes and validation checks before data enters One Model and regular monitoring and updating of data sources to ensure that the data reflects current conditions.

And finally, we return to regulatory compliance. One AI's configurability supports adherence to various regulatory requirements while its fully auditable nature ensures that all model activities are traceable, thus supporting compliance efforts. As we continue to develop the tool, we maintain a forward looking approach to regulatory compliance, ensuring that our tool stays ahead of evolving requirements. Best practices around regulatory compliance include staying updated with the latest regulatory changes and incorporating them into your organization's machine learning governance framework.

Chapter 4

Ethical ML is Good for Business

Section 4 - Ethical Machine Learning is Good for Business

We designed One AI to be fully transparent and ethical, not only because it's the right thing to do, but also because it makes good business sense. This approach mitigates legal and compliance risks, fosters trust, enhances decision making, and ensures fair treatment for people, ultimately benefiting both organizations and their workforce.

Adhering to ethical standards and privacy laws, such as GDPR or CCPA reduces the risk of legal penalties and compliance issues. Ethical machine learning practices help organizations meet regulatory requirements, avoiding costly fines and damage to reputation.

Ethical machine learning practices build trust among employees, stakeholders, and the public. Trust is crucial for the successful adoption and effective use of people analytics.

When stakeholders trust that the analytics are conducted responsibly, they are more likely to support and engage with the initiatives, enhancing overall business operations.

Conversely, unethical practices can lead to public scandals, customer backlash, and long term reputational damage.

Ethical machine learning ensures decisions are made based on accurate, fair, and unbiased data. This can improve the quality of decisions related to hiring promotions and other HR functions, which improves overall organizational performance.

And finally, ethical practices identify and eliminate biases in algorithms preventing discriminatory outcomes. This fair treatment enhances inclusion within the organization and is key in attracting and retaining top talent.

Chapter 5

Conclusion & Thanks

As we conclude this module, it's clear that ethical machine learning is crucial for sustainable and successful business practices.

One AI was carefully designed with ethics in mind, ensuring more accurate, fair, and transparent decision making. By adopting the best practices we've discussed, your organization can protect its reputation, ensure compliance, and drive positive change in the workforce. Remember, ethical machine learning is an ongoing journey that requires continuous commitment and vigilance. Happy modeling!