In this module, you will:

* Describe various bias mitigating algorithms that can be intervened in a ML Pipeline
* Explain the usage of various bias mitigating algorithms
* Recognize the role of open source toolkit AI Fairness 360 in mitigating bias
* Describe various features and capabilities of opensource AI Fairness toolkit
* Analyze the working of opensource AI Fairness toolkit with an interactive demo

To be successful in this module, prior knowledge is recommended in :

* Module 1: The big picture of Trustworthy AI and Algorithmic Fairness
* Data Science / Machine Learning Workflow
* Evaluation metrics

Bias occurs in data used to train a model. In this interactive demo, three sample datasets have been provided that you can use to explore bias checking and mitigation. Each dataset contains attributes that should be protected to avoid bias. Take some time to explore the outputs and mitigation strategies.

Click <https://aif360.mybluemix.net/data> link to open resource.

Demos, Tutorials, APIs and Additional Guidance

This extensible open source toolkit can help you examine, report, and mitigate discrimination and bias in machine learning models throughout the AI application lifecycle. We invite you to use and improve it.

Click <https://aif360.mybluemix.net/> link to open resource.

#### AI Fairness 360 Community

Connect with AI Fairness researchers, experts and developers

* Link to the slack community: <https://aif360.slack.com/>

#### O'Reilly Book - AI Fairness: How to Measure and Reduce Unwanted Bias in Machine Learning

This report takes data science leaders and practitioners through the key challenges of defining fairness and reducing unfair bias throughout the machine learning pipeline. It shows why data science teams need to engage early and authoritatively on building trusted artificial intelligence (AI). It also explains in plain English how organizations should think about AI fairness, as well as the trade-offs between model bias and model accuracy. Much has been written on the social injustice of AI bias, but this report focuses on how teams can mitigate unfair machine bias by using the open source tools available in AI Fairness 360 (AIF360).

* Link to the report: <http://ibm.biz/Bdqbd2>

#### Research Paper on AI Fairness 360

Fairness is an increasingly important concern as machine learning models are used to support decision making in high-stakes applications such as mortgage lending, hiring, and prison sentencing. This article introduces a new open-source Python toolkit for algorithmic fairness, AI Fairness 360.

Link to the report: [Research Paper on AI Fairness 360 (20min) - PDF](https://learn.ibm.com/pluginfile.php/1317415/mod_page/content/3/Research%20Paper%20on%20AI%20Fairness%20360%20%2820min%29%20-%20PDF)

Provided below is a summary of the entire module. For your convenience, below you will see a heading that matches the content topics that were covered in this module, followed by a bulleted list of key concepts covered.

***Lesson 1: Mitigating Bias in ML Pipeline***

1. Bias can occur at any stage in the machine learning pipeline. The following **bias mitigation algorithms** can be performed at various stages of your pipeline depending on your relevancy.

* **Pre-processing**: Reduce bias by manipulating the training data prior to training the algorithm.
* **In-processing**: Reduce bias by creating fairness constraints as input and use them to influence the loss in the training algorithm
* **Post-processing:** Reduce bias by manipulating the training data after training the algorithm

2. IBM researchers have created the **Open Source toolkit AI Fairness 360**, which is a comprehensive Toolkit for Detecting, Understanding, and Mitigating Unfair Bias.

3. There are tradeoffs between bias & accuracy i.e reducing bias may lower your model accuracy.

***Lesson 2: Overview of Opensource Toolkit AI Fairness 360***

4. The AI Fairness 360 [interactive demo](https://aif360.mybluemix.net/data?_ga=2.120519179.2012469676.1618842117-2008629828.1615927664&cm_mc_uid=99344246712216159276575&cm_mc_sid_50200000=81148061619200602501) provides a gentle introduction to the concepts and capabilities. This [website](https://aif360.mybluemix.net/?_ga=2.120519179.2012469676.1618842117-2008629828.1615927664&cm_mc_uid=99344246712216159276575&cm_mc_sid_50200000=81148061619200602501) gives you an overview of this extensible open source toolkit. The [complete API](https://aif360.readthedocs.io/) is also available.