

Fairness in Machine Learning Models for Banking Applications

Machine learning models have become increasingly popular in the banking industry, particularly in credit risk, financial crime, and fraud detection. However, these models can inadvertently perpetuate and even amplify biases, leading to unfair and discriminatory outcomes. Fairness in machine learning models for banking applications is an important consideration to ensure that lending and other financial services are provided equitably to all individuals or groups, without discrimination based on protected characteristics such as disability, gender, or age.

At Barclays, ensuring fairness in machine learning models has become a routine activity in our model development pipeline. This talk will cover various techniques to identify and mitigate potential sources of bias in our data, such as exploring the distribution of the data by protected characteristics and testing for mathematical approaches such as demographic parity or equal opportunity metrics. We will also discuss how we leverage explainability methods to understand how the model makes decisions and detect potential biases. Finally, we will detail how we perform regular audits and monitoring of our models in production to ensure ongoing fairness and compliance with regulatory requirements. Overall, by incorporating fairness into our model development pipeline, we can ensure that our machine learning models for banking applications are not only accurate and effective, but also fair and ethical.