Algorithmic foundations and ethics in AI: from theory to practice course

Toolkit for synchronous sessions

CU3 | Algorithms and their limitations

Case study

Project number | 2022-1-ES01-KA220-HED-000085257

Case study

Case study
Description for trainers

	Description
Task description	Students will analyze a fictional case (see next slide) involving a financial technology (FinTech) company that used an algorithmic system to assess creditworthiness for loan approvals. The company aimed to provide quick and efficient loan services to a broader demographic. However, it was later discovered that the algorithm consistently favored applicants from certain income brackets and racial backgrounds, resulting in discriminatory loan practices. Students must investigate the root causes of this issue, evaluate the consequences on affected communities, and suggest effective mitigation strategies.
Description of how to do the task	Group work 3-4 persons group. **Research & Background Analysis** • Review the course materials and literature on algorithmic bias and fairness. • Understand the specific factors that contributed to FinTech Innovations' discriminatory outcomes, including data biases, features used, and the absence of fairness metrics. **Root Cause Identification** • Analyze the ways in which FinTech Innovations' training data and features could have embedded societal biases. • Investigate the absence of fairness metrics and the lack of stakeholder input. **Mitigation Recommendations** Develop a set of recommendations to address the identified biases, including data diversification, new validation procedures, ethical guidelines, and ongoing auditing. **Drafting & Submission** Organize findings in a structured report, providing a clear summary, analysis of root causes, and detailed mitigation strategies. Include visuals, graphs, or tables where relevant to support your points.
Estimated time to do the task	30-60 minutes.
Suggestion of sources for doing the task	Course materials, academic databases, GAI systems (i.e. ChatGPT, Copilot, etc.) and internet searches.
Detailed description of how to deliver the task	Write a 500-750 words summary report.
Information on the deadline for the task delivery	During the synchronous session.
Contact information or how to clarify doubts	The teacher must provide a form of contact (it could be an email address, a telephone number...)

Case study
Fictional case for students

The FinTech industry is known for its innovative approach to providing quick and efficient financial services, often through the use of algorithmic systems for assessing creditworthiness and approving loans. This fictional case revolves around a FinTech company that employed an advanced credit scoring algorithm to evaluate applicants and streamline its loan approval process. The company's mission was to democratize access to financial resources and support customers who are often underserved by traditional banks. However, its algorithm unexpectedly resulted in discriminatory lending practices, systematically favoring applicants from certain income brackets and racial backgrounds.

Company Background

The fictional FinTech company, FinTech Innovations, was established to disrupt the traditional banking sector by offering a seamless, fully online loan approval system. This system utilized a proprietary algorithm to assess credit risk based on various factors, such as income, employment history, geographic location, and educational background. FinTech Innovations aimed to tap into markets ignored by mainstream banks, particularly among younger individuals, minority groups, and those with unconventional income sources.

Problem Identification

Despite its stated goals of inclusivity, FinTech Innovations' credit scoring algorithm consistently favored applicants from higher income brackets and specific racial backgrounds. This resulted in disproportionately high loan rejection rates for lower-income applicants and those from racial minorities. Discrepancies in loan approval rates across demographic groups raised concerns among customer advocacy groups, leading to an external audit that confirmed the algorithm's discriminatory behavior.

Root Cause Analysis

The analysis revealed several root causes of the bias. First, the historical data used to train the algorithm reflected existing societal disparities, embedding systemic biases into the model. For instance, it used geographic location as a feature, which correlated strongly with income and race. The algorithm further relied heavily on traditional employment data, overlooking those with unconventional income streams like freelancers or gig workers. Finally, the lack of fairness metrics during model validation allowed these biases to persist undetected.

Impact Assessment

The discriminatory practices of FinTech Innovations had significant social and financial repercussions for affected groups. Loan rejection rates for minorities and low-income applicants were nearly double those of higher-income applicants from privileged backgrounds. This exclusion resulted in missed opportunities for financial growth and further entrenched existing inequalities. The company faced reputational damage, regulatory scrutiny, and legal challenges.

Case study
Fictional case for students

Mitigation Strategies

To address these issues, FinTech Innovations implemented several corrective measures:

- **Data Diversification**
  - They broadened the training dataset to include diverse applicants from different income levels, races, and employment types.
- **Fairness Metrics**
  - They incorporated fairness metrics to ensure the algorithm's predictions were unbiased and representative.
- **Transparency Measures**
  - The company committed to sharing more information on how its algorithm assessed creditworthiness.
- **Ethical Guidelines**
  - New guidelines required continuous monitoring, impact assessment, and stakeholder engagement.

Conclusion and Recommendations

This case illustrates the unintended consequences of algorithmic decision-making in FinTech. It underscores the importance of using diverse training data, validating algorithms with fairness metrics, and maintaining transparency to mitigate bias and discrimination. Industry-wide ethical standards and regulatory frameworks are essential for fostering fair and inclusive FinTech practices.

THANK YOU

Project number | 2022-1-ES01-KA220-HED-000085257

Co-funded by the European Union