Aggregato alla
FACOLTÀ DI SCIENZE DELL'EDUCAZIONE - UNIVERSITÀ PONTIFICIA SALESIANA

Toward a cost effective fairness-aware Machine Learning Lifecycle

Ital - IA 2023

Bussi Elisa[1], Basso Andrea[1], Maran Elena[2],
Chiavarino Claudia[1], Severini Simone[2]

(1) Istituto Universitario Salesiano Torino Rebaudengo (ius.to)
(2) Modulos AG (modulos.ai)

Introduction

AI/ML adoption is set to change the way we do many things, including **making decisions**.

This comes along with several concerns about the possibility of AI to amplify existing **social biases**, following the occurrence of some discrimination episodes:

- **Robert Williams**
- **Amazon's HR AI**
- **Joey Buolamwuini**

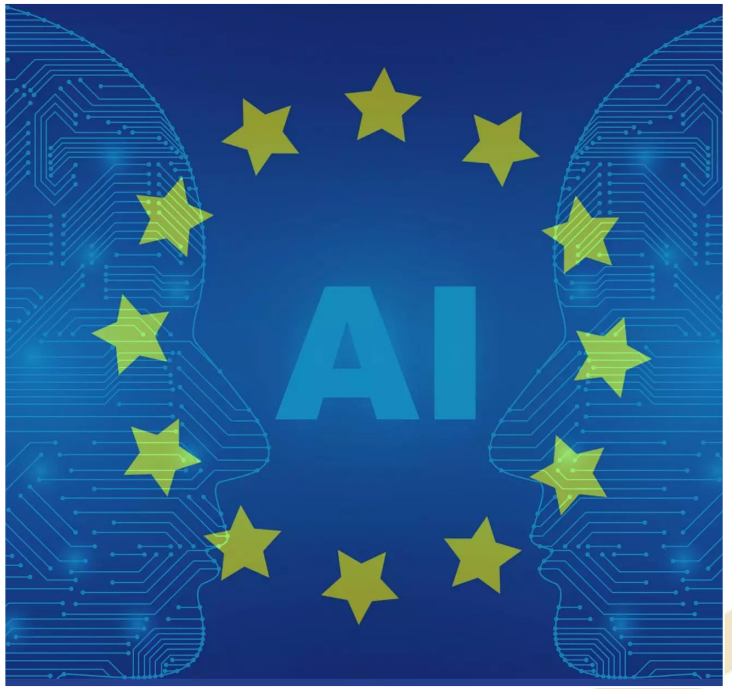
Joy Buolamwini, pioneer activist that advocates for algorithmic justice. https://www.ajl.org/

The EU AI Act

GDPR and **EU AI Act** will make **considering fairness** in AI/ML that has an impact on people **mandatory**.

Aside for compliance, being fair has many advantages:

- **reduced risk of failure**
- **enhanced user trust**
- **reinforced brand reputation**

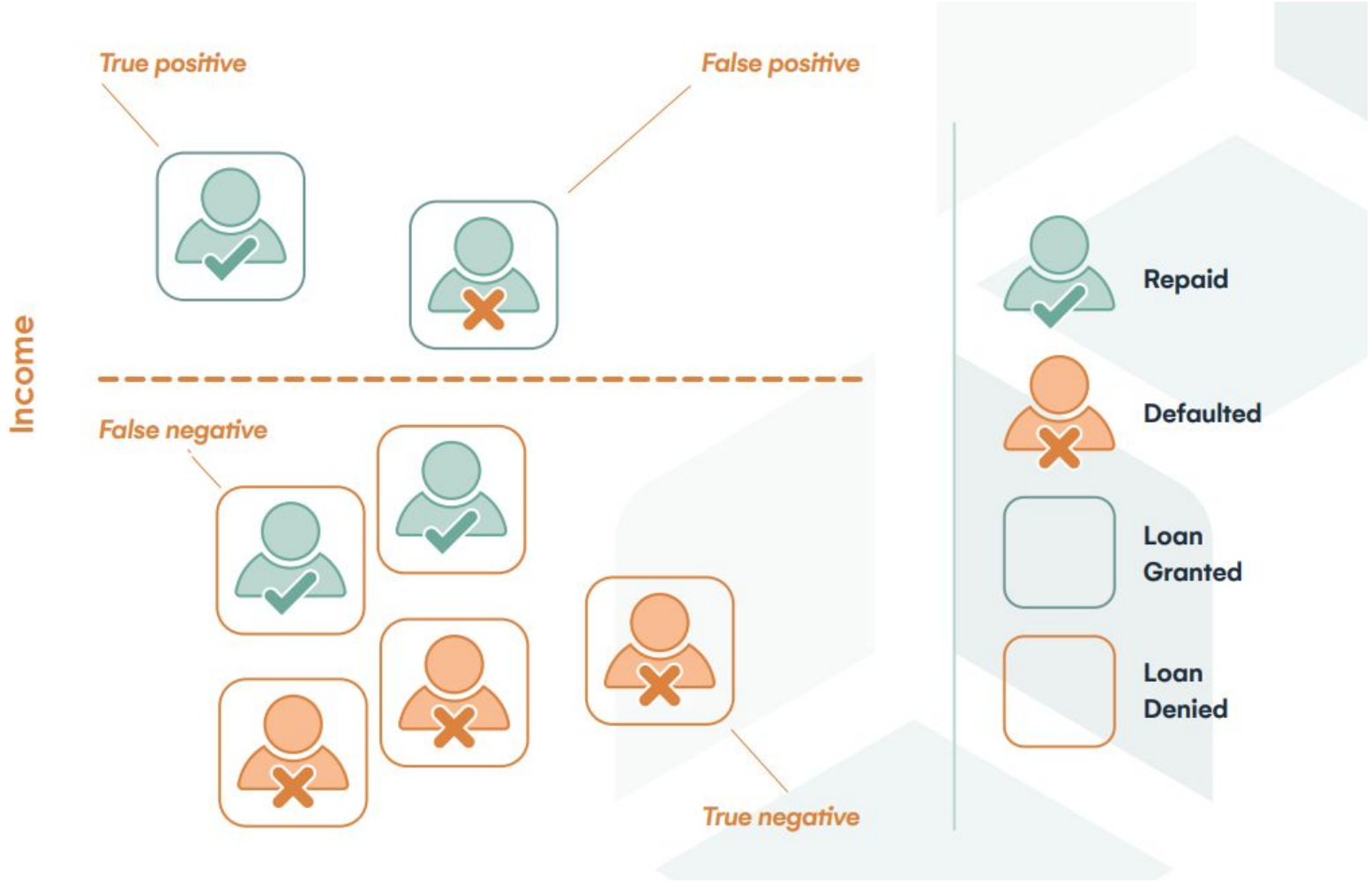


Fairness Integration: Practitioner Difficulties

ML practitioners struggle with implementing fairness in everyday practice because:

- **Unclear** how to properly **comply with regulations**
- **No systematic procedure** for fairness implementation
- **Fairness tools not optimized** for practitioners' workflow
- **Lack** of **fairness tools' coverage** of ML Lifecycle
- **Organizational constraints/barrier** against fairness implementation

Example

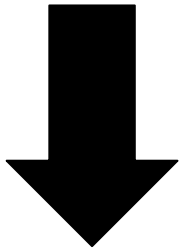


Exploring AI Fairness in Consumer Lending - Modulos

How do we decide how to make a **fair prediction system** in order to grant a loan based on certain eligibility criteria?

Fairness highly depends on the context!

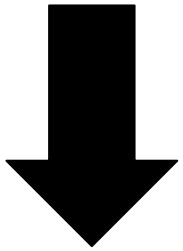
Want to achieve **equal treatment** for each sub-group?



Demographic Parity

historical bias reinforcement

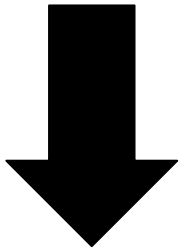
Want to **maximize** the amount of **predictions made correctly**?



Equal Opportunity

critical harm in case of **misclassification**

Want to 'play it safe' and **minimize losses**?



Predictive Parity

frequent **discard** of possible **profitable opportunities**

Current approaches

Z-Inspection®

IUS.TO Survey

What is the state of the art of **fairness implementation** in the **italian AI ecosystem**?

Survey Focus:

- **Knowledge** about Ethical and Fair AI/ML
- **Systemacy** of Fairness throughout ML Lifecycle
- Fairness tools **usage** and **usability**
- **Individual/Organizational** barriers towards Fairness

Purpose: understand how to **overcome challenges** in **ML Fairness implementation** and how to **support practitioners** in transitioning towards a **fairness-aware AI**.

Thank you for your attention… and the help you will give us!