A Low-Code Approach for Machine Learning Fairness Applied to IoT Data in Health Domain

Giordano d'Aloisio, Antinisca Di Marco

Università degli Studi dell'Aquila / Italy

giordano.daloisio@graduate.univaq.it

Outline

- Introduction
- Proposed Approach
- Use Case
- Live Demo

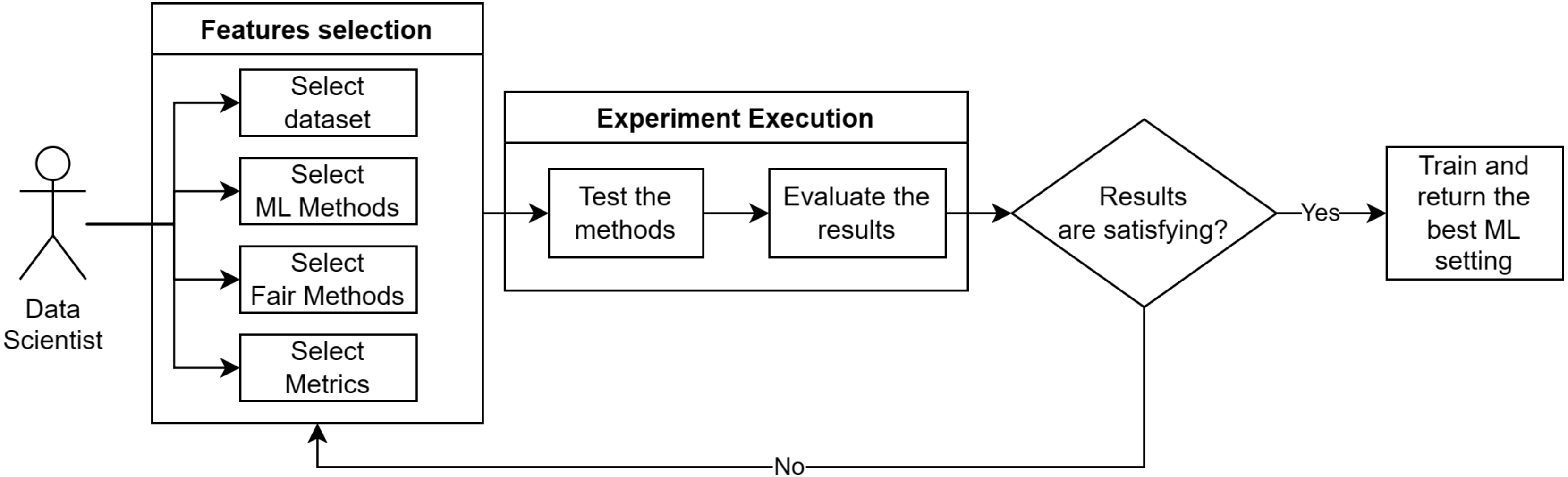
Introduction

- Machine Learning systems are widely used nowadays in many domains
- Many low-code systems have been proposed to make the ML development accessible also to non-expert users

- However, a low-code tool covering essential quality properties (e.g., fairness, explainability, privacy) of ML systems is still missing
- In this work, we target the Fairness quality property by presenting MANILA a low code platform for development of fair ML systems

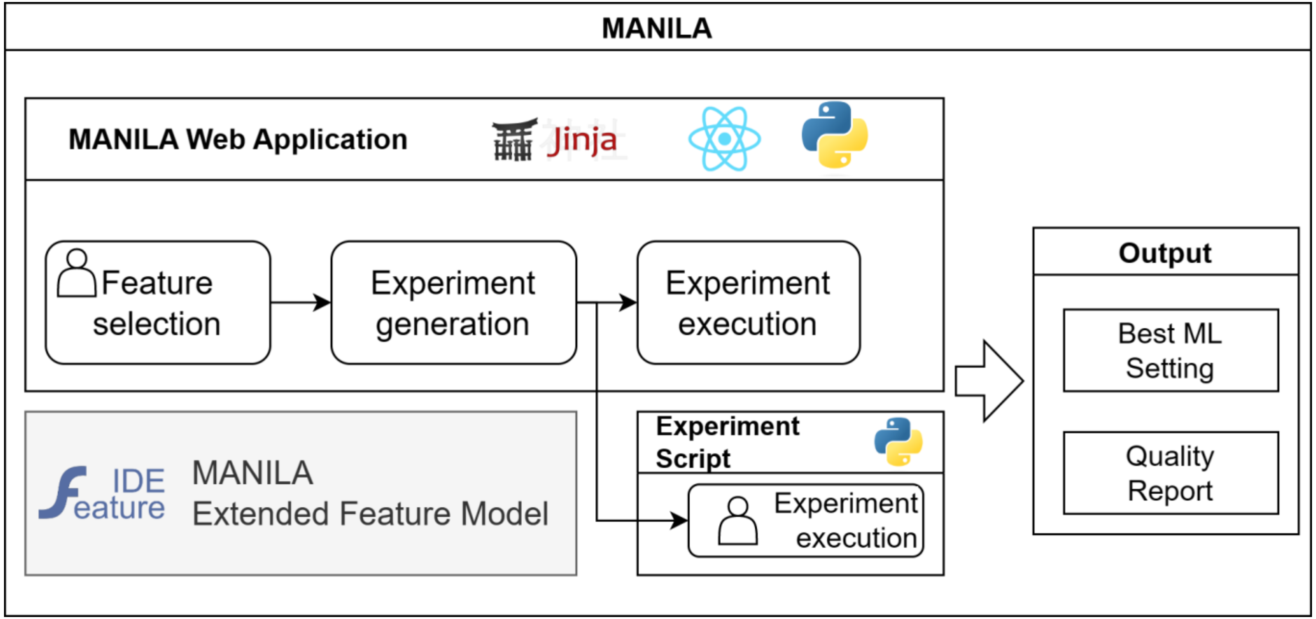
Fairness development workflow

- A general workflow for the development of fair ML systems can be sketched as following



MANILA [1]

- MANILA is a low-code tool to support the development of fair ML systems



[1] Giordano d'Aloisio, Antinisca Di Marco, Giovanni Stilo "Democratizing Quality-Based Machine Learning Development through Extended Feature Models", FASE 2023

Fairness

- Fairness is the absence of prejudice or favoritism by a ML model over items identified by a set of *sensitive variables*
- Fairness is usually defined by a set of relevant concepts

Sensitive variables

Privileged and unprivileged groups

Favorable outcome

Fairness Metrics

- Measuring fairness means measuring the amount of discrimination performed by an ML model
- Metrics like Disparate Impact or Equal Opportunity measure the probability of getting the positive outcome predicted being in the unprivileged group or not

Nowadays is not possible to measure a priori the fairness of an ML model

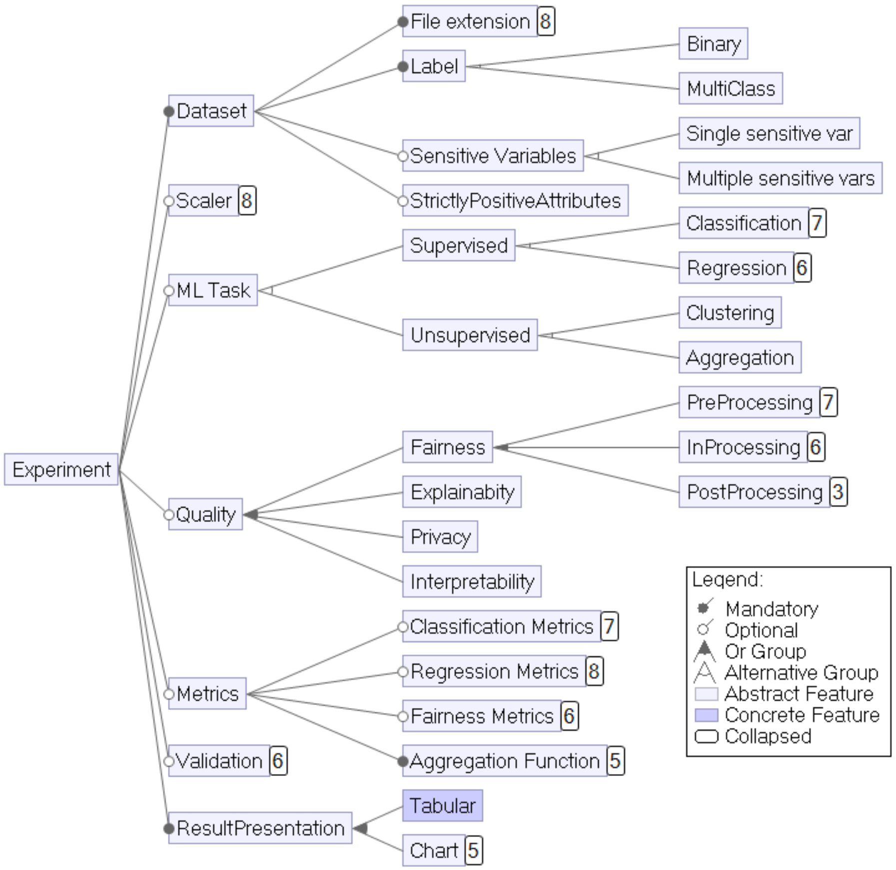
Fairness Enhancing Methods

- Fairness enhancing approaches can be classified in three categories

Pre-processing methods	**Work on the dataset to reduce its bias**
In-processing methods	**Change the training process to make it fair**
Post-processing methods	**Work on an already trained model to remove the learned bias**

MANILA's Extended Feature Model

- MANILA is based on an Extended Feature Model that models each feature of an experimental evaluation



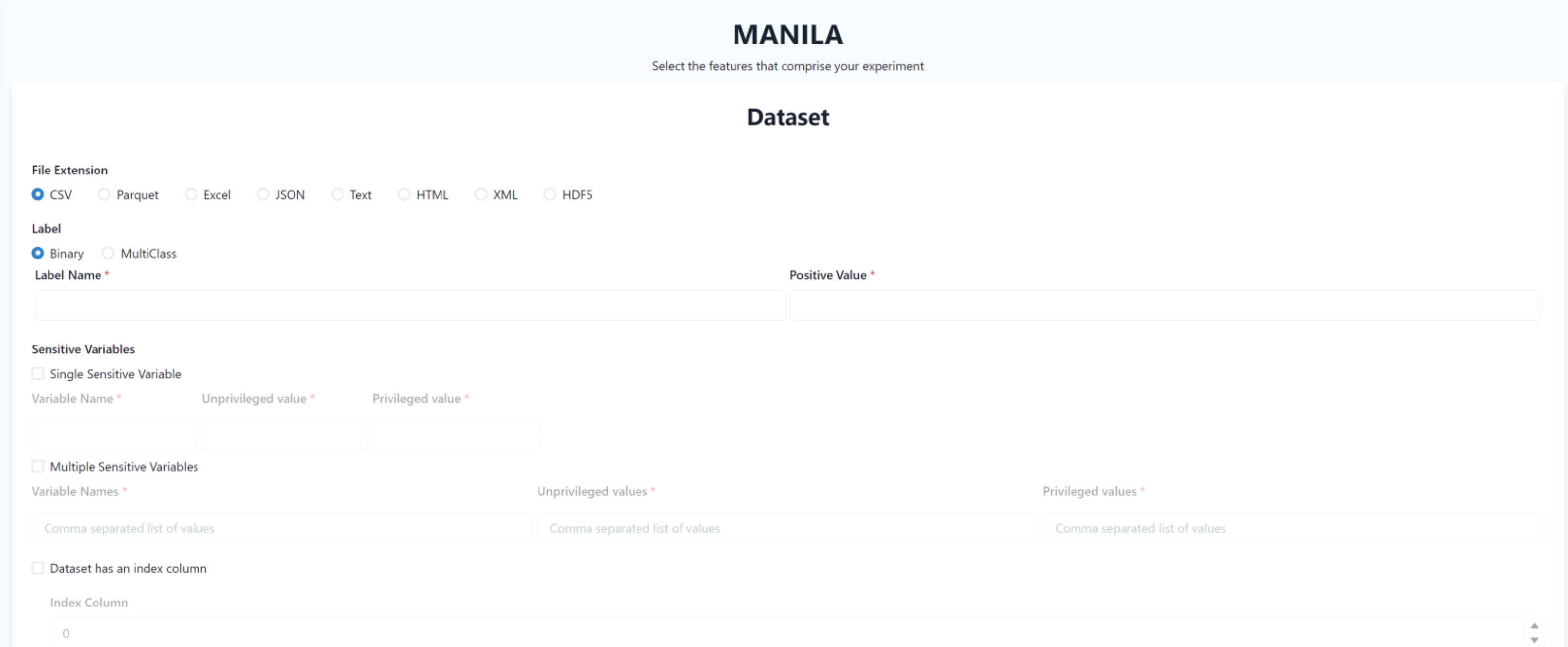
Extended Feature Model Constraints

- These constraints guide the data scientist in the selection of features that always lead to a correct experiment

Fairness ⇒ "Sensitive Variables"
ExponentiatedGradient ∨ GridSearch ⇒ ¬"MLP Classifier" ∧ ¬"MLP Regressor"
¬GerryFairClassifier ∧ ¬MetaFairClassifier ∧ ¬AdversarialDebiasing ⇒ "ML Task"
Classification ⇔ "Classification Metrics" ∧ ¬"Regression Metrics"
"Classification Metrics" ⇒ ¬"Regression Metrics"
Regression ⇔ "Regression Metrics" ∧ ¬"Classification Metrics"
Fairness ⇒ "Fairness Metrics"
"BoxCox Method" ⇒ StrictlyPositiveAttributes
¬DIR ∨ ¬"Multiple sensitive vars"
¬MultiClass ∨ ¬Reweighing
¬DIR ∨ ¬MultiClass
¬AdversarialDebiasing ∨ ¬MultiClass
¬MultiClass ∨ ¬GerryFairClassifier
¬MultiClass ∨ ¬MetaFairClassifier
¬MultiClass ∨ ¬PrejudiceRemover
¬MultiClass ∨ ¬CalibratedEO
¬MultiClass ∨ ¬RejectOptionClassifier
"Multiple sensitive vars" ⇒ ¬PostProcessing
SVC ⇒ ¬PostProcessing
Regression ⇒ ¬Fairness
"Gradient Descent Classifier" ⇒ ¬PostProcessing
Reweighing ⇒ ¬"MLP Classifier"
AUC ⇒ ¬MultiClass

Extended Feature Model Implementation

- The Extended Feature Model has been implemented as a web form with a set of constraints among the fields



Use Case Scenario

SCG-RHC Dataset

- Dataset of 73 patients to monitor changes in pulmonary arterial pressure (PAP) and pulmonary capillary wedge pressure (PCWP) using a chest-worn wearable patch.
- Provides
  - Demographic information of the patients
  - Signals in the dataset recorded by the wearable patch
- Useful to predict the clinical status of patient
- https://physionet.org/content/scg-rhc-wearable-database/1.0.0/

Experiment Description

- Train a Logistic Regression model to predict if a patient is under *physiological challenge*
- Analyze if the model is biased against *women*, if so, mitigate the bias

ML Model	Fairness Method	Metrics	Aggregation Function
Logistic Regression	DEMV	Accuracy Equalized Odds	Harmonic Mean

MANILA in Action