

Module 4 Lab:

Classifier Audit with FairVis

OBJECTIVE

Machine learning has promoted the development of artificial intelligence significantly. However, during the process of build machine learning models, it will introduce implicit and explicit biases to these models, which will generate misleading outputs from these models. Discovering such biases is a great challenging task, due to the numerous definitions of fairness and the large number of potentially impacted subgroups. This lab is to leverage FairVis (<https://github.com/poloclub/FairVis>), a mixed-initiative visual analytics system that integrates a novel subgroup discovery technique for users to audit the fairness of machine learning models, to discover model biases on the classifier for the breast cancer classification.

PREREQUISITES

You have to figure out this paper (<https://arxiv.org/abs/1904.05419>) about FairVis and its repository (<https://github.com/poloclub/FairVis>). In addition, you must install FairVis and figure out how to run it with demo.

INSTRUCTIONS

- Step 1: Download the breast cancer data (breast-cancer-data.csv) from UCI data repository (<https://archive.ics.uci.edu/ml/datasets/Breast+Cancer>).
- Step 2: Insert this row (Class,age,menopause,tumor-size,inv-nodes,node-caps,deg-malig,breast,breast-quad,breast-quad) as the data header on the top of the data file (breast-cancer-data.csv).
- Step 3: Build a classifier with **Logistic Regression** on the data and create a new file (breast_out.csv) with the last two columns being the output class of the model and the ground truth labels. Examples of models in Jupyter Notebook format can be found in ./models of the FairVis repository (<https://github.com/poloclub/FairVis>).
- Step 4: Preprocess the classified data (breast_out.csv) by running **Kmean** clustering with 4 ($K = 4$) clusters and save the processed file (breast.csv) to ./src/data/ of the FairVis repository.
- Step 5: Import the processed file (breast.csv) in the src/components/Welcome.js component of the FairVis repository by adding a new row to the table in Welcome.js around line 188 in the form of the other datasets.
- Step 6: Create a subgroup with features **Age** and **Breast**.
- Step 7: Discovery model biases by comparing performance differences between (age (30-39) + breast (right)) (Pinned) and (age (60-69) + breast (right)) (Hovered), where the performance will be presented with Accuracy, Precision, and Recall.