XAI_HW1

We have two populations Blue (privileged) and Red (unprivileged), with the Blue population being 9 times larger than the Red population.

Individuals from both populations are requesting to attend XAI training to improve competency in this important area. Number of places is limited. The administrators of the training have decided to give priority to enrolling individuals who may need this training in the future, although unfortunately it is difficult to predict who will benefit.

The decision rule adopted:

1. In the Red group, half of the people will find the skills useful in future and half will not. Administrators randomly allocate 50% of people to training.
2. in the Blue group, 80% of people will find the training useful in future and 20% will not, although of course it is not known who will find it useful. The administrators have built a predictive model based on user behaviour in predicting for whom it will be useful and whom will not. The model has the following performance:

Blue	Will use XAI	Will not use XAI	Total
Enrolled in training	60	5	65
not enrolled in training	20	15	35
Total	80	20	100

Task: Calculate the Demographic parity, equal opportunity and predictive rate parity coefficients for this decision rule.

Starred task: How can this decision rule be changed to improve its fairness?

To work on integer numbers we can assume following numbers of
$Total\ RED = 100$
$Total\ BLUE = 900$

Demographic parity- *measures how decisions are dependent to being in specific group, in our case there are only two groups*

- $P(X\ enrolled|X \in RED) = 0.5$
- $P(X\ enrolled|X \in BLUE) = \frac{enrolled}{total} = 0.65$
- $DP = \frac{0.5}{0.65} = 0.769$ lub $|0.65 - 0.5| = 0.15$ (czasami liczone jako różnica między najwyższym a najniższym wynikiem pośród grup)

Equal opportunity

- The ratio of people who enrolled and will use XAI to total number who will use XAI.
  $P(X\ enrolled|\ will\ use \wedge RED) = \frac{enrolled \wedge will\ use}{will\ use\ total} = \frac{0.5 \times 0.5}{0.5} = 0.5$
  $P(X\ enrolled|\ will\ use \wedge BLUE) = \frac{60}{80} = 0.75$
  $EO = |0.5 - 0.75| = 0.25$

Predictive rate parity coefficients

- $P(will\ use|enrolled \wedge BLUE) = \frac{60}{65} = 0.923$

- $P(will\ use|enrolled \wedge RED) = \frac{0.5 \times 0.5}{0.5} = 0.5$
- $PRT = 0.923 - 0.5 = 0.423$