

Non-discriminatie by design

Wat is het probleem en hoe gaat AI helpen het probleem op te lossen? Welk percentage in foutnegatieven en foutpositieven is acceptabel?

1

Probleemdefinitie

2

Dataverzameling

In hoeverre zijn de benodigde data al binnen de organisatie beschikbaar en in hoeverre moeten ze van buiten worden gehaald? Welke bias zit er in de data?

3

Datavoorbereiding

Aan de hand van welke criteria wordt de keuze gemaakt om bepaalde data wel en andere niet te gebruiken voor het model en hoe raakt deze keuze aan onderscheid tussen groepen?

Modellering

4

Hoe worden criteria op het gebied van uitlegbaarheid en fairness vertaald naar de modelselectiestrategie?
Hoe presteert het model op de gekozen benchmark van foutpositieven en foutnegatieven?

6

Evaluatie

Kies een implementatiestrategie en formuleer een evaluatie- en exitstrategie. Beoordeel hoe het systeem functioneert en hoe het zou functioneren met een ander model, fairness definitie en/of algoritme.

5

Implementatie

Selecteer een afgebakende en representatieve toepassing om het systeem te testen. Pas het model aan op basis van de resultaten en betrek belanghebbenden.

