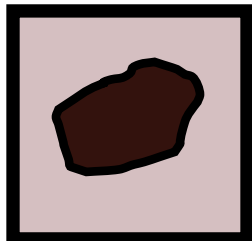


Original Image



AI Classifier

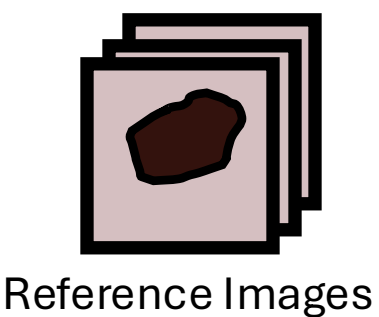


Protected Attribute
Prediction

Sex	Prob
Male	0.8
Female	0.2

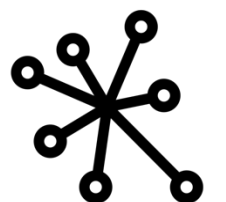
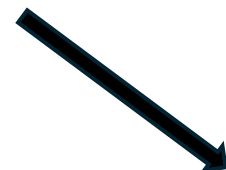


Which signals are important?



+

AI Classifier



Clustering

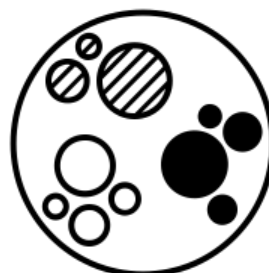


Image Clusters

+

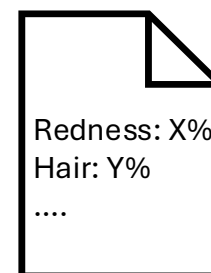


Signals Identified



Propensity Score
Reweighting

Signals Quantified



Generative AI

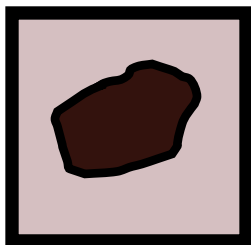


Counterfactuals

+



Original Image



AI Classifier



Protected Attribute
Prediction

Sex	Prob
Male	0.8
Female	0.2



Which signals are important?



Statistical
Associations

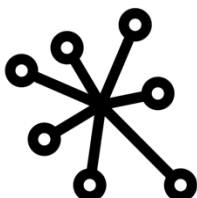


Disease	Attribute
a	b
c	d

Odds Ratio



Reference Images
+



Clustering



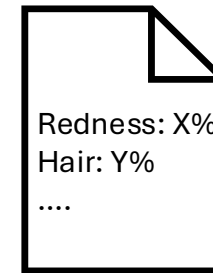
Image Clusters

+



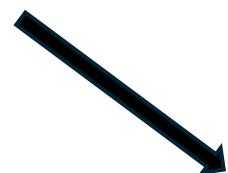
Signals Identified

Propensity Score
Reweighting



Signals Quantified

AI Classifier



Generative AI



Counterfactuals

+

