Einleitung

Der Bereich der Künstlichen Intelligenz (KI) ist ein rasant wachsender Forschungszweig, mit unzähligen wertvollen Beiträgen aus der ganzen Welt. Dabei sind die Anwendungsmöglichkeiten genauso vielfältig wie die verschiedenen Datenquellen dazu. Da diese Datenquellen allerdings alle direkt oder indirekt durch Menschen erschaffen sind, repräsentieren sie teilweise auch deren Vorurteile und Weltbilder. Auch wenn diese Verzerrungen größtenteils unbeabsichtigt in die Daten projiziert werden, sind sie dennoch vorhanden und können dazu führen, dass sich die KI unfair verhält und bestimmten Gruppen Vorteile gegenüber anderen Gruppen verschafft. So kann es beispielsweise vorkommen, dass weißen Männern eher ein Kredit gewährt wird als Frauen oder dunkelhäutigen Menschen, wenn für diese Entscheidung ein Machine Learning (ML) Algorithmus verwendet wird. Unser Ziel dabei ist es, ein Artefakt zu entwickeln, welches solche Vorurteile/Verzerrungen (Bias) in einem tabellarischen Datensatz erkennt und ihn idealerweise gleich bereinigt.