Replication steps

- Replace model and dataset paths appropriately (same directory as scripts)
- Default settings used (COMPAS dataset + model):
    - Number of test cases: 250
    - Sensitive attribute: "race"
    - Perturbation probability: 0.3
    - Perturbation noise level: 0.05
    - Improved tool threshold: 0.05 (for flagging bias based on prediction score difference)
- Default settings used (KDD dataset + model):
    - Improved tool threshold: 0.04 (for flagging bias based on prediction score difference)
    - All other parameters the same as for COMPAS
- Changes required to make code word for KDD dataset + model:
    - In the COMPAS dataset, the race column is “Race” (capitalised), whereas in the KDD dataset it is “race”
    - Therefore all mentions of the column heading in the code need to be replaced with lowercase “race”
    - The *target_column* attribute is also different for the KDD dataset. It needs to be changed from “Recidivism” to “income” (case sensitive). This is the column that is dropped before evaluation of the model so that only input features are considered.