**Section 3 and 4** of the “tutorial\_medical\_expenditure.ipynb” notebook discusses Random Forest (RF) classifier as an alternative classifier that can be used for predicting healthcare utilization, with and without bias mitigation.

1. Run the **subsections** **3.3** (learning RF classifier on original data) and **4.3** (learning RF classifier on data transformed by reweighing) and compare the results with the corresponding ones for Logistic Regression (LR) classifier. How effective is reweighing preprocessing on RF compared to LR classifier with respect to best balanced accuracy, 1-min(DI, 1/DI) and average odds difference measures?
2. Rerun the cells in **Section 4.3** five times without restarting the notebook every time. Observe and record what happens to the best balanced accuracy, 1-min(DI, 1/DI) and average odds difference measures on test data over the five trials.
3. Set the hyperparameters n\_estimators=50 and min\_samples\_leaf=2 when creating the RF pipeline in **Section 4.3.1**.

Now repeat question ‘B’ in this document

With respect to the three measures, compare the stability of the original hyperparameter setting to the new hyperparameter setting. If we were to deploy this bias-mitigated classifier on a new test set, which  hyperparameter setting should be used, and why?