Preprocessing

Testing Data
holdout.py

Training & Validation Data

Impute or Drop NAs*
ML_preprocess.py

↓

Feature Selection*
Feature_Selection.py

↓

Features to use

*random balanced draws **X10***

Grid Search

10-fold cross-validation

Training & Validation Data

Train Parameter Sweep

↓

Test Parameter Sweep

↓

Parameters to use

*random balanced draws **X100***

Traing & Validation

10-fold cross-validation

Training & Validation Data

Train Model

↓

Validate Model

↓

Performance Statistics

Apply to Testing & Unknown

Train Model

Apply

Testing Data

Unknown Data

↓

Performance Statistics

↓

Predict Class/Y

***Optional**