Adversarial Attack

Method of Application

Evaluation

Robustness Predictive

Accuracy

Predictions (A1)

Predictions

Predictions (A2)

Robustness Explanations

Explanations

Explanations (A1)

Euclidean

Distance

AUC

AUC

, comparison

comparison



