Likelihood Ratio Attack Report

Introduction

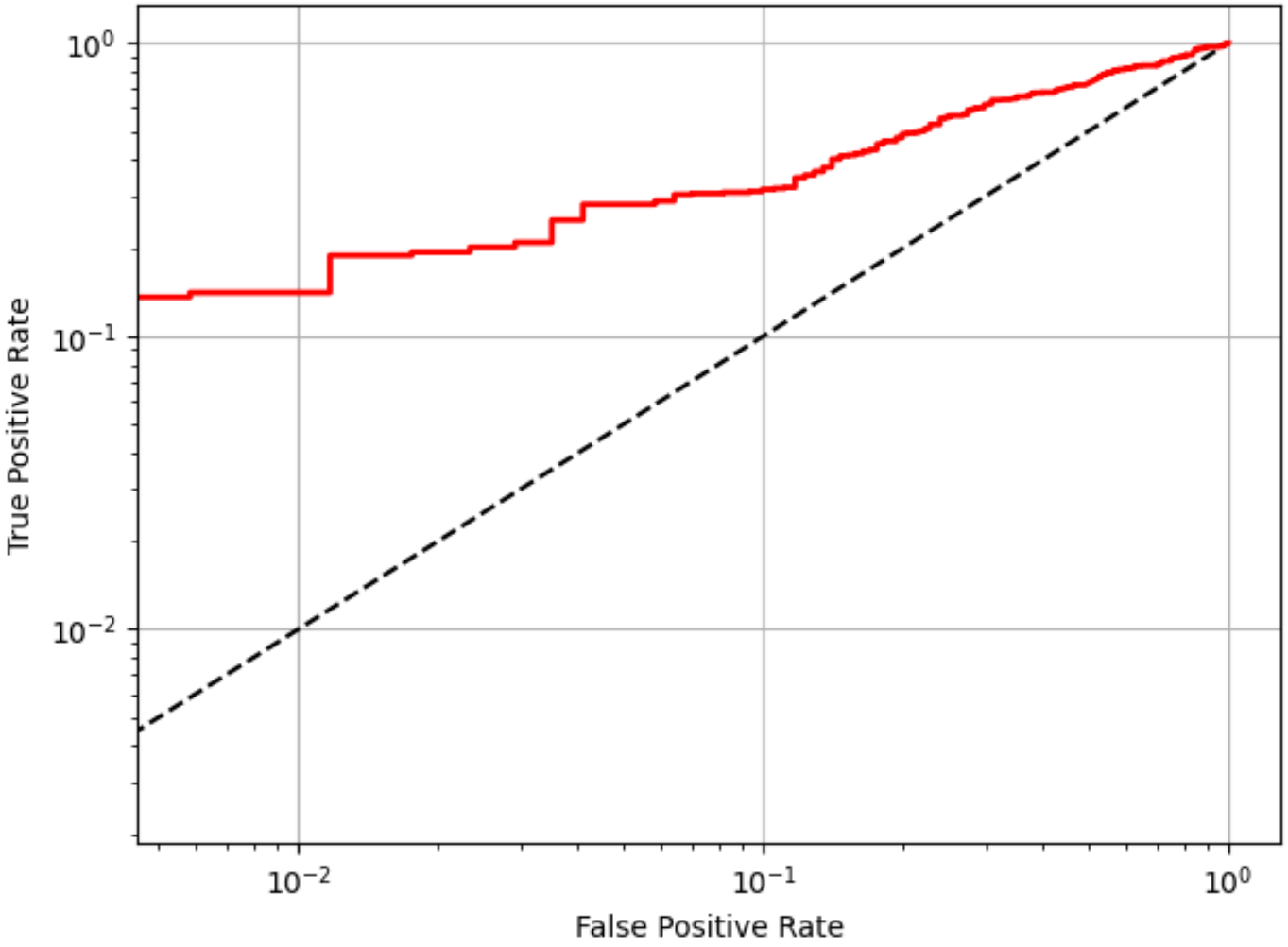
Metadata

```
            n_shadow_models: 100
                   p_thresh: 0.05
                 output_dir: outputs_lira
                report_name: report_lira
     training_data_filename: train_data.csv
         test_data_filename: test_data.csv
    training_preds_filename: train_preds.csv
        test_preds_filename: test_preds.csv
               target_model: ['sklearn.ensemble', 'RandomForestClassifier']
           target_model_hyp: {'min_samples_split': 2, 'min_samples_leaf': 1}
 attack_config_json_file_name: lira_config.json
n_shadow_rows_confidences_min: 10
      shadow_models_fail_fast: False
                  target_path: None
                         mode: offline
                fix_variance: False
           report_individual: False
                    PDIF_sig: Significant at p=0.05
                     AUC_sig: Significant at p=0.05
         null_auc_3sd_range: 0.4207446718814698 -> 0.5792553281185302
```

Metrics

```
                       TPR: 0.7186
                       FPR: 0.4912
                       FAR: 0.2270
                       TNR: 0.5088
                       PPV: 0.7730
                       NPV: 0.4372
                       FNR: 0.2814
                       ACC: 0.6555
                   F1score: 0.7448
                 Advantage: 0.2274
                       AUC: 0.6957
              P_HIGHER_AUC: 0.0000
                    FMAX01: 0.9825
                    FMIN01: 0.5088
                    FDIF01: 0.4737
                    PDIF01: 0.0000
                    FMAX02: 0.9386
                    FMIN02: 0.5614
                    FDIF02: 0.3772
                    PDIF02: 22.0591
                   FMAX001: 1.0000
                   FMIN001: 0.5000
                   FDIF001: 0.5000
                   PDIF001: 3.5249
             pred_prob_var: 0.0849
                  n_normal: 0.7083
```

ROC Curve


10^0
10^{-1}
10^{-2}
True Positive Rate
10^{-2}
10^{-1}
10^0
False Positive Rate

Likelihood Ratio Attack Report

Introduction

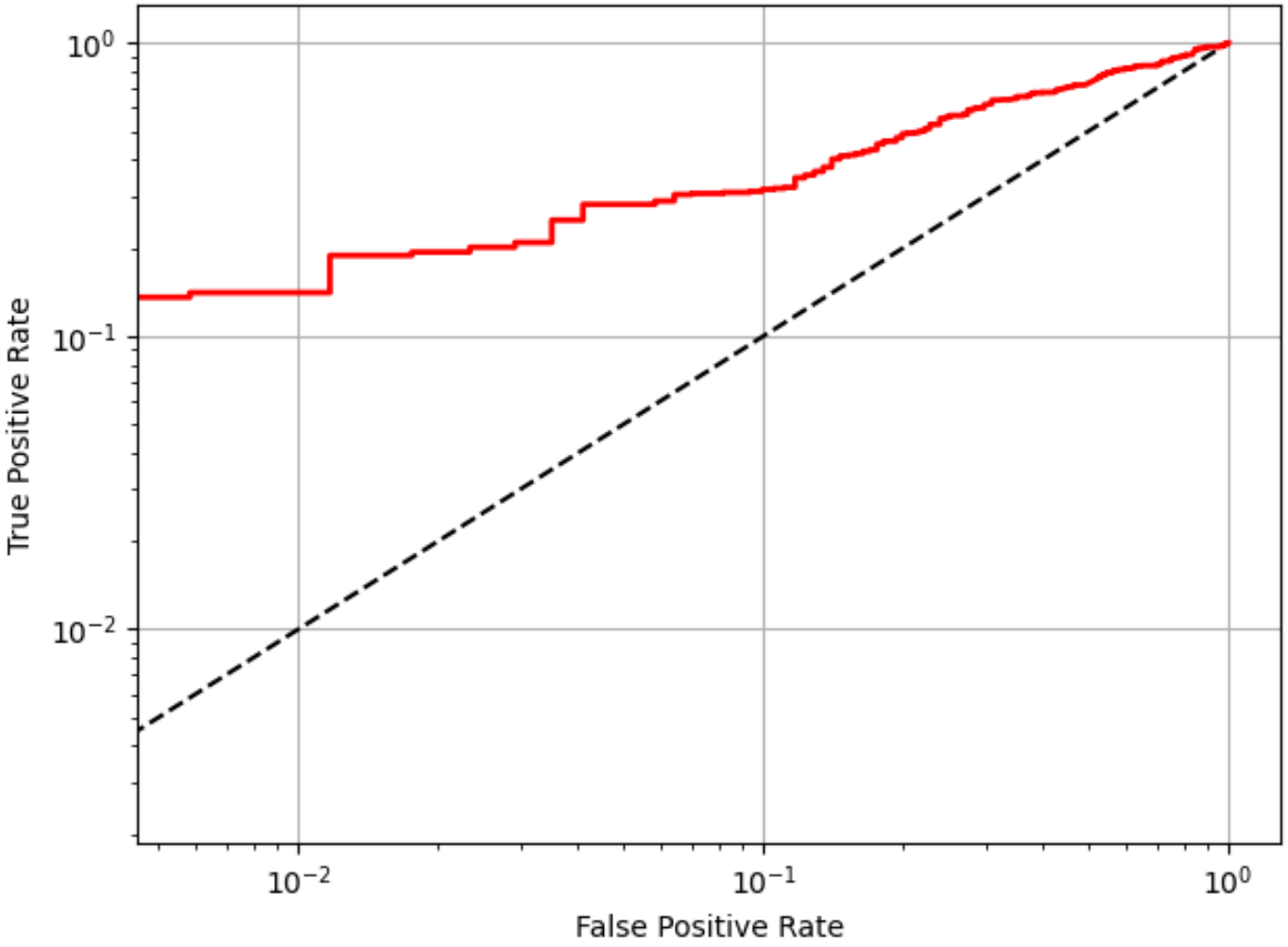
Metadata

```
            n_shadow_models: 100
                   p_thresh: 0.05
                 output_dir: outputs_lira
                report_name: report_lira
     training_data_filename: train_data.csv
         test_data_filename: test_data.csv
    training_preds_filename: train_preds.csv
        test_preds_filename: test_preds.csv
               target_model: ['sklearn.ensemble', 'RandomForestClassifier']
           target_model_hyp: {'min_samples_split': 2, 'min_samples_leaf': 1}
 attack_config_json_file_name: lira_config.json
n_shadow_rows_confidences_min: 10
      shadow_models_fail_fast: True
                  target_path: None
                         mode: offline
                fix_variance: False
           report_individual: False
                    PDIF_sig: Significant at p=0.05
                     AUC_sig: Significant at p=0.05
         null_auc_3sd_range: 0.4207446718814698 -> 0.5792553281185302
```

Metrics

```
                       TPR: 0.7186
                       FPR: 0.4912
                       FAR: 0.2270
                       TNR: 0.5088
                       PPV: 0.7730
                       NPV: 0.4372
                       FNR: 0.2814
                       ACC: 0.6555
                   F1score: 0.7448
                 Advantage: 0.2274
                       AUC: 0.6957
              P_HIGHER_AUC: 0.0000
                    FMAX01: 0.9825
                    FMIN01: 0.5088
                    FDIF01: 0.4737
                    PDIF01: 0.0000
                    FMAX02: 0.9386
                    FMIN02: 0.5614
                    FDIF02: 0.3772
                    PDIF02: 22.0591
                   FMAX001: 1.0000
                   FMIN001: 0.5000
                   FDIF001: 0.5000
                   PDIF001: 3.5249
             pred_prob_var: 0.0849
                  n_normal: 0.7083
```

ROC Curve


10^0
10^{-1}
10^{-2}
True Positive Rate
10^{-2}
10^{-1}
10^0
False Positive Rate

Likelihood Ratio Attack Report

Introduction

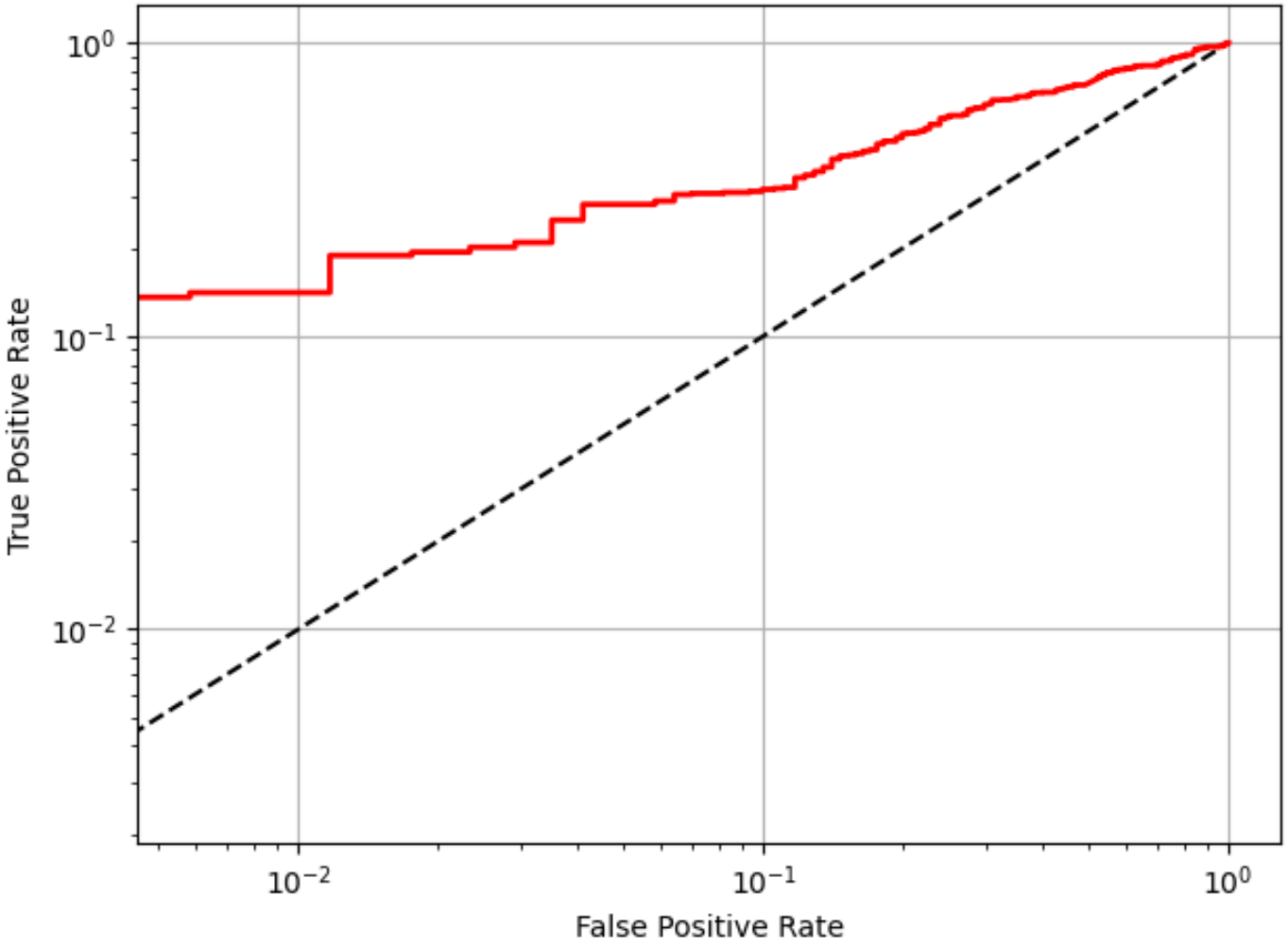
Metadata

```
            n_shadow_models: 100
                   p_thresh: 0.05
                 output_dir: outputs_lira
                report_name: report_lira
     training_data_filename: train_data.csv
         test_data_filename: test_data.csv
    training_preds_filename: train_preds.csv
        test_preds_filename: test_preds.csv
               target_model: ['sklearn.ensemble', 'RandomForestClassifier']
           target_model_hyp: {'min_samples_split': 2, 'min_samples_leaf': 1}
 attack_config_json_file_name: lira_config.json
n_shadow_rows_confidences_min: 10
      shadow_models_fail_fast: False
                  target_path: None
                         mode: offline
                fix_variance: False
           report_individual: False
                    PDIF_sig: Significant at p=0.05
                     AUC_sig: Significant at p=0.05
         null_auc_3sd_range: 0.4207446718814698 -> 0.5792553281185302
```

Metrics

```
                       TPR: 0.7186
                       FPR: 0.4912
                       FAR: 0.2270
                       TNR: 0.5088
                       PPV: 0.7730
                       NPV: 0.4372
                       FNR: 0.2814
                       ACC: 0.6555
                   F1score: 0.7448
                 Advantage: 0.2274
                       AUC: 0.6957
              P_HIGHER_AUC: 0.0000
                    FMAX01: 0.9825
                    FMIN01: 0.5088
                    FDIF01: 0.4737
                    PDIF01: 0.0000
                    FMAX02: 0.9386
                    FMIN02: 0.5614
                    FDIF02: 0.3772
                    PDIF02: 22.0591
                   FMAX001: 1.0000
                   FMIN001: 0.5000
                   FDIF001: 0.5000
                   PDIF001: 3.5249
             pred_prob_var: 0.0849
                  n_normal: 0.7083
```

ROC Curve


10^0
10^{-1}
10^{-2}
True Positive Rate
10^{-2}
10^{-1}
10^0
False Positive Rate

Likelihood Ratio Attack Report

Introduction

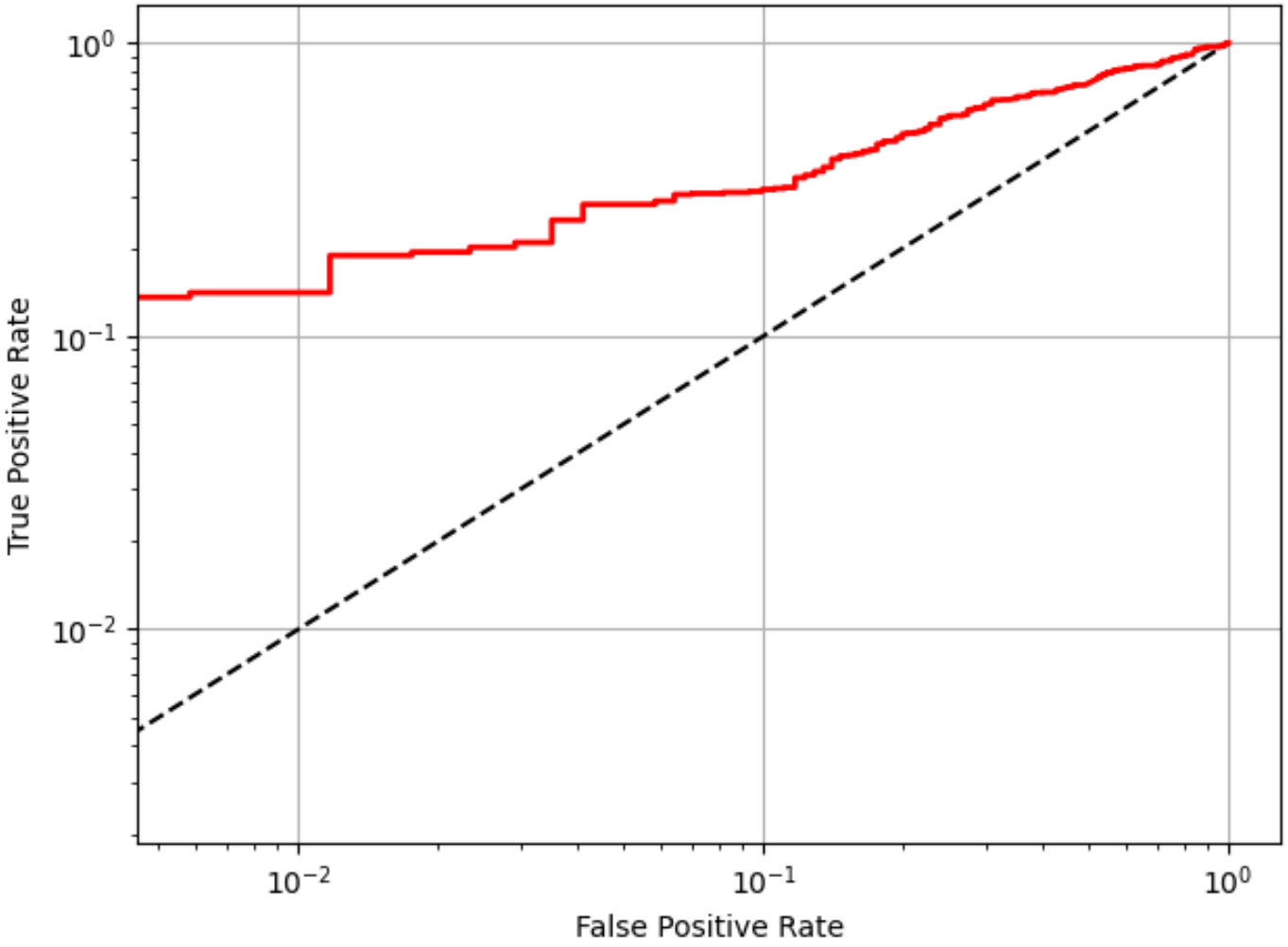
Metadata

```
            n_shadow_models: 100
                   p_thresh: 0.05
                 output_dir: outputs_lira
                report_name: report_lira
     training_data_filename: train_data.csv
         test_data_filename: test_data.csv
    training_preds_filename: train_preds.csv
        test_preds_filename: test_preds.csv
               target_model: ['sklearn.ensemble', 'RandomForestClassifier']
           target_model_hyp: {'min_samples_split': 2, 'min_samples_leaf': 1}
 attack_config_json_file_name: lira_config.json
n_shadow_rows_confidences_min: 10
      shadow_models_fail_fast: True
                  target_path: None
                         mode: offline
                fix_variance: False
           report_individual: False
                    PDIF_sig: Significant at p=0.05
                     AUC_sig: Significant at p=0.05
         null_auc_3sd_range: 0.4207446718814698 -> 0.5792553281185302
```

Metrics

```
                       TPR: 0.7186
                       FPR: 0.4912
                       FAR: 0.2270
                       TNR: 0.5088
                       PPV: 0.7730
                       NPV: 0.4372
                       FNR: 0.2814
                       ACC: 0.6555
                   F1score: 0.7448
                 Advantage: 0.2274
                       AUC: 0.6957
              P_HIGHER_AUC: 0.0000
                    FMAX01: 0.9825
                    FMIN01: 0.5088
                    FDIF01: 0.4737
                    PDIF01: 0.0000
                    FMAX02: 0.9386
                    FMIN02: 0.5614
                    FDIF02: 0.3772
                    PDIF02: 22.0591
                   FMAX001: 1.0000
                   FMIN001: 0.5000
                   FDIF001: 0.5000
                   PDIF001: 3.5249
             pred_prob_var: 0.0849
                  n_normal: 0.7083
```

ROC Curve


10^0
10^{-1}
10^{-2}
True Positive Rate
10^{-2}
10^{-1}
10^0
False Positive Rate

Likelihood Ratio Attack Report

Introduction

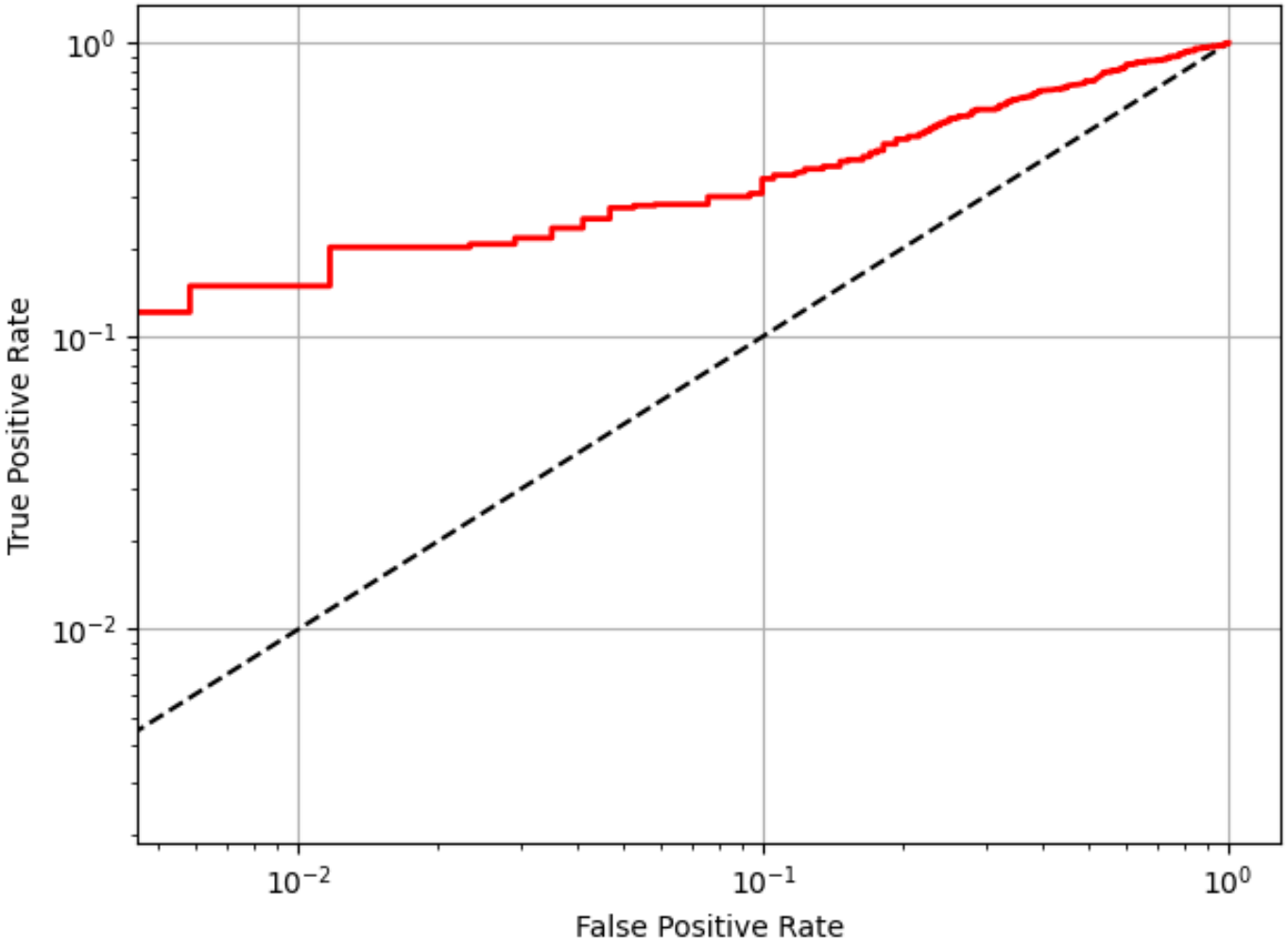
Metadata

```
            n_shadow_models: 150
                   p_thresh: 0.05
                 output_dir: outputs_lira
                report_name: report_lira
     training_data_filename: train_data.csv
         test_data_filename: test_data.csv
    training_preds_filename: train_preds.csv
        test_preds_filename: test_preds.csv
               target_model: ['sklearn.ensemble', 'RandomForestClassifier']
           target_model_hyp: {'min_samples_split': 2, 'min_samples_leaf': 1}
 attack_config_json_file_name: config_lira_cmd1.json
n_shadow_rows_confidences_min: 10
      shadow_models_fail_fast: False
                  target_path: target_model_for_lira
                         mode: offline
                fix_variance: False
           report_individual: False
                    PDIF_sig: Significant at p=0.05
                     AUC_sig: Significant at p=0.05
          null_auc_3sd_range: 0.4207446718814698 -> 0.5792553281185302
```

Metrics

```
                       TPR: 0.7462
                       FPR: 0.5146
                       FAR: 0.2286
                       TNR: 0.4854
                       PPV: 0.7714
                       NPV: 0.4511
                       FNR: 0.2538
                       ACC: 0.6678
                   F1score: 0.7586
                 Advantage: 0.2316
                       AUC: 0.6993
              P_HIGHER_AUC: 0.0000
                    FMAX01: 0.9825
                    FMIN01: 0.4561
                    FDIF01: 0.5263
                    PDIF01: 0.0000
                    FMAX02: 0.9298
                    FMIN02: 0.4825
                    FDIF02: 0.4474
                    PDIF02: 30.0686
                   FMAX001: 1.0000
                   FMIN001: 0.6667
                   FDIF001: 0.3333
                   PDIF001: 2.2637
             pred_prob_var: 0.0793
                  n_normal: 0.7575
```

ROC Curve


True Positive Rate
False Positive Rate
10^{0}
10^{-1}
10^{-2}
10^{-2}
10^{-1}
10^{0}

Likelihood Ratio Attack Report

Introduction

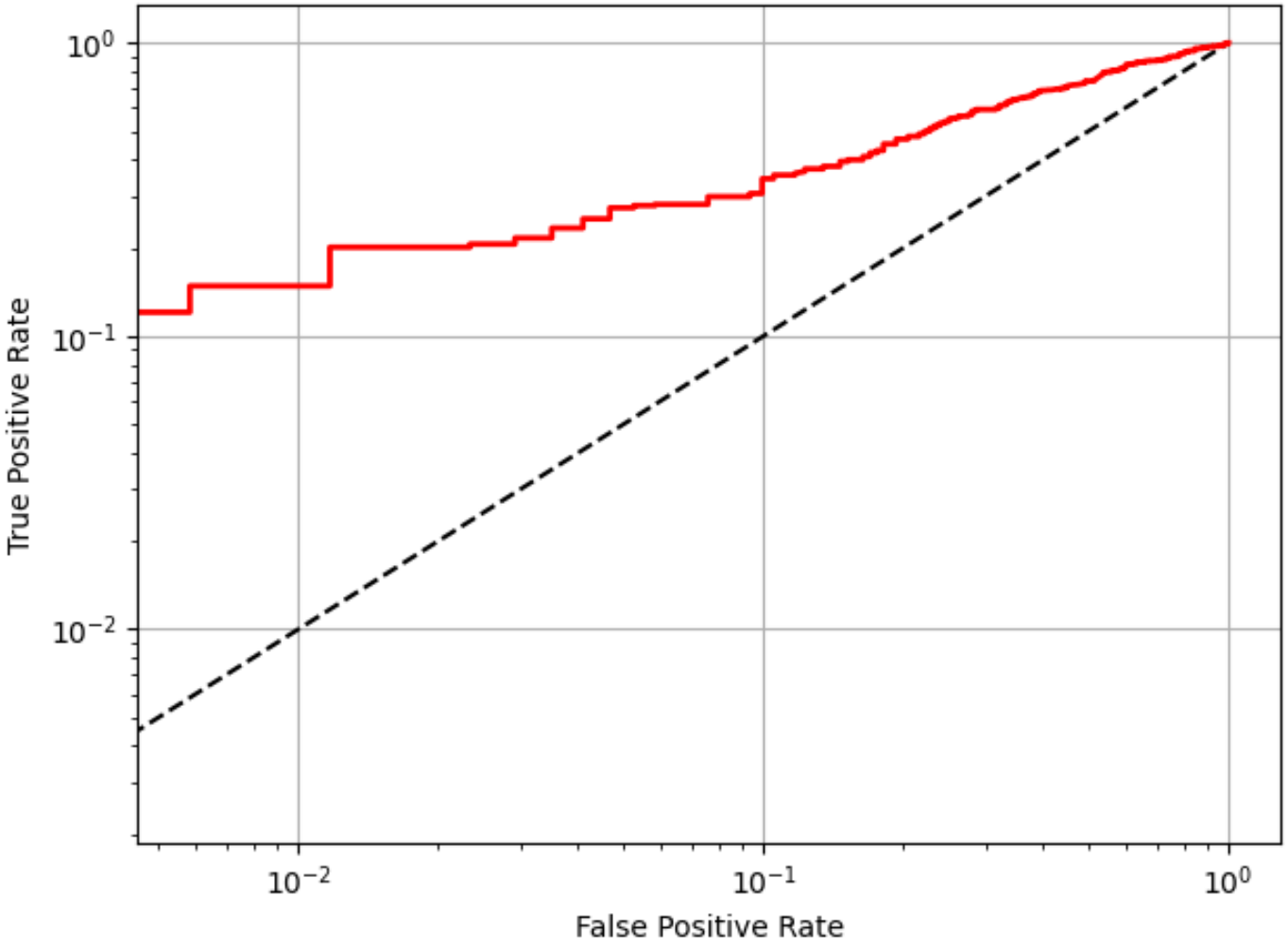
Metadata

```
            n_shadow_models: 150
                   p_thresh: 0.05
                 output_dir: outputs_lira
                report_name: report_lira
     training_data_filename: train_data.csv
         test_data_filename: test_data.csv
    training_preds_filename: train_preds.csv
        test_preds_filename: test_preds.csv
               target_model: ['sklearn.ensemble', 'RandomForestClassifier']
           target_model_hyp: {'min_samples_split': 2, 'min_samples_leaf': 1}
 attack_config_json_file_name: config_lira_cmd2.json
n_shadow_rows_confidences_min: 10
      shadow_models_fail_fast: True
                  target_path: target_model_for_lira
                         mode: offline
                fix_variance: False
           report_individual: False
                    PDIF_sig: Significant at p=0.05
                     AUC_sig: Significant at p=0.05
          null_auc_3sd_range: 0.4207446718814698 -> 0.5792553281185302
```

Metrics

```
                       TPR: 0.7462
                       FPR: 0.5146
                       FAR: 0.2286
                       TNR: 0.4854
                       PPV: 0.7714
                       NPV: 0.4511
                       FNR: 0.2538
                       ACC: 0.6678
                   F1score: 0.7586
                 Advantage: 0.2316
                       AUC: 0.6993
              P_HIGHER_AUC: 0.0000
                    FMAX01: 0.9825
                    FMIN01: 0.4561
                    FDIF01: 0.5263
                    PDIF01: 0.0000
                    FMAX02: 0.9298
                    FMIN02: 0.4825
                    FDIF02: 0.4474
                    PDIF02: 30.0686
                   FMAX001: 1.0000
                   FMIN001: 0.6667
                   FDIF001: 0.3333
                   PDIF001: 2.2637
             pred_prob_var: 0.0793
                  n_normal: 0.7575
```

ROC Curve


True Positive Rate
False Positive Rate
10^{0}
10^{-1}
10^{-2}
10^{-2}
10^{-1}
10^{0}