

1 Duplicates

1.1 AllFeatures

Table 1: Results for AllFeatures on Duplicates Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9452	0.0132	0.9403	0.0143	0.9366	0.0208	0.9384	0.0151
RandomForest	0.9610	0.0111	0.9788	0.0098	0.9329	0.0221	0.9551	0.0131
GradientBoosting	0.9610	0.0109	0.9663	0.0066	0.9455	0.0199	0.9557	0.0127
XGBoost	0.9610	0.0150	0.9676	0.0163	0.9442	0.0209	0.9557	0.0170

1.2 Intersection1

Table 2: Results for Intersection1 on Duplicates Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9491	0.0050	0.9413	0.0181	0.9455	0.0153	0.9431	0.0052
RandomForest	0.9638	0.0070	0.9764	0.0064	0.9417	0.0171	0.9586	0.0083
GradientBoosting	0.9621	0.0119	0.9642	0.0167	0.9506	0.0157	0.9573	0.0134
XGBoost	0.9644	0.0102	0.9691	0.0140	0.9506	0.0146	0.9597	0.0115

1.3 Intersection2

Table 3: Results for Intersection2 on Duplicates Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9299	0.0142	0.9267	0.0284	0.9164	0.0216	0.9211	0.0159
RandomForest	0.9525	0.0120	0.9707	0.0138	0.9214	0.0185	0.9454	0.0140
GradientBoosting	0.9497	0.0145	0.9535	0.0164	0.9329	0.0217	0.9430	0.0166
XGBoost	0.9491	0.0189	0.9569	0.0196	0.9278	0.0292	0.9420	0.0219

1.4 KeywordFeatures

Table 4: Results for KeywordFeatures on Duplicates Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9237	0.0158	0.9310	0.0256	0.8961	0.0221	0.9129	0.0175
RandomForest	0.9446	0.0148	0.9741	0.0167	0.8999	0.0255	0.9353	0.0177
GradientBoosting	0.9395	0.0147	0.9538	0.0211	0.9088	0.0235	0.9305	0.0170
XGBoost	0.9384	0.0125	0.9648	0.0179	0.8948	0.0210	0.9283	0.0149

1.5 ManualSelection

Table 5: Results for ManualSelection on Duplicates Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9316	0.0093	0.9305	0.0089	0.9151	0.0194	0.9226	0.0112
RandomForest	0.9491	0.0121	0.9667	0.0079	0.9177	0.0282	0.9413	0.0148
GradientBoosting	0.9514	0.0112	0.9560	0.0131	0.9341	0.0162	0.9449	0.0128
XGBoost	0.9514	0.0090	0.9682	0.0110	0.9214	0.0202	0.9441	0.0108

1.6 MutualInformation1

Table 6: Results for MutualInformation1 on Duplicates Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9519	0.0062	0.9504	0.0126	0.9417	0.0147	0.9459	0.0071
RandomForest	0.9650	0.0111	0.9739	0.0124	0.9468	0.0153	0.9601	0.0127
GradientBoosting	0.9627	0.0142	0.9665	0.0173	0.9493	0.0183	0.9578	0.0161
XGBoost	0.9610	0.0142	0.9675	0.0136	0.9442	0.0209	0.9557	0.0163

1.7 MutualInformation2

Table 7: Results for MutualInformation2 on Duplicates Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9344	0.0098	0.9279	0.0137	0.9253	0.0227	0.9263	0.0114
RandomForest	0.9525	0.0125	0.9694	0.0130	0.9227	0.0205	0.9454	0.0146
GradientBoosting	0.9525	0.0119	0.9598	0.0170	0.9329	0.0176	0.9460	0.0134
XGBoost	0.9508	0.0087	0.9647	0.0164	0.9240	0.0207	0.9436	0.0102

1.8 ParkingSensores

Table 8: Results for ParkingSensores on Duplicates Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9062	0.0116	0.8901	0.0274	0.9025	0.0248	0.8957	0.0123
RandomForest	0.9339	0.0090	0.9447	0.0136	0.9050	0.0199	0.9242	0.0107
GradientBoosting	0.9333	0.0116	0.9289	0.0177	0.9214	0.0221	0.9249	0.0134
XGBoost	0.9322	0.0101	0.9285	0.0119	0.9189	0.0213	0.9235	0.0120

1.9 RFE1

Table 9: Results for RFE1 on Duplicates Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9457	0.0082	0.9408	0.0177	0.9379	0.0185	0.9391	0.0092
RandomForest	0.9633	0.0129	0.9751	0.0133	0.9417	0.0221	0.9580	0.0150
GradientBoosting	0.9616	0.0172	0.9641	0.0215	0.9493	0.0188	0.9566	0.0193
XGBoost	0.9638	0.0115	0.9715	0.0119	0.9468	0.0199	0.9589	0.0133

1.10 RFE2

Table 10: Results for RFE2 on Duplicates Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9356	0.0103	0.9381	0.0156	0.9164	0.0213	0.9269	0.0120
RandomForest	0.9553	0.0117	0.9709	0.0121	0.9278	0.0221	0.9487	0.0138
GradientBoosting	0.9508	0.0147	0.9619	0.0152	0.9265	0.0238	0.9438	0.0171
XGBoost	0.9548	0.0131	0.9647	0.0165	0.9328	0.0158	0.9485	0.0149

2 Simple

2.1 AllFeatures

Table 11: Results for AllFeatures on Simple Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9208	0.0286	0.7574	0.1105	0.7670	0.0875	0.7580	0.0831
RandomForest	0.9454	0.0156	0.9148	0.0758	0.7327	0.0744	0.8102	0.0570
GradientBoosting	0.9427	0.0193	0.8718	0.0823	0.7556	0.0645	0.8083	0.0648
XGBoost	0.9390	0.0250	0.8726	0.1107	0.7330	0.0856	0.7939	0.0838

2.2 Intersection1

Table 12: Results for Intersection1 on Simple Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9272	0.0190	0.7970	0.0975	0.7559	0.0606	0.7703	0.0493
RandomForest	0.9418	0.0176	0.8640	0.0760	0.7614	0.0733	0.8068	0.0597
GradientBoosting	0.9399	0.0185	0.8453	0.0698	0.7668	0.0561	0.8036	0.0598
XGBoost	0.9390	0.0205	0.8366	0.0801	0.7722	0.0611	0.8024	0.0660

2.3 Intersection2

Table 13: Results for Intersection2 on Simple Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.8990	0.0353	0.7045	0.1219	0.6816	0.0565	0.6888	0.0826
RandomForest	0.9272	0.0135	0.8450	0.0828	0.6760	0.0394	0.7490	0.0432
GradientBoosting	0.9208	0.0182	0.8279	0.1168	0.6594	0.0327	0.7296	0.0480
XGBoost	0.9245	0.0131	0.8323	0.0881	0.6763	0.0407	0.7425	0.0344

2.4 KeywordFeatures

Table 14: Results for KeywordFeatures on Simple Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.8926	0.0146	0.7026	0.0772	0.5965	0.0757	0.6390	0.0488
RandomForest	0.9235	0.0134	0.8826	0.0800	0.6083	0.0424	0.7186	0.0458
GradientBoosting	0.9181	0.0210	0.8208	0.1233	0.6483	0.0514	0.7193	0.0610
XGBoost	0.9217	0.0161	0.8235	0.0973	0.6651	0.0560	0.7320	0.0507

2.5 ManualSelection

Table 15: Results for ManualSelection on Simple Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9035	0.0181	0.7209	0.0753	0.6597	0.0551	0.6872	0.0544
RandomForest	0.9263	0.0157	0.8393	0.0861	0.6757	0.0413	0.7467	0.0483
GradientBoosting	0.9272	0.0126	0.8293	0.0729	0.6932	0.0212	0.7540	0.0370
XGBoost	0.9308	0.0186	0.8420	0.0875	0.7044	0.0393	0.7665	0.0585

2.6 MutualInformation1

Table 16: Results for MutualInformation1 on Simple Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9245	0.0281	0.7631	0.1023	0.7841	0.0858	0.7701	0.0808
RandomForest	0.9472	0.0160	0.9122	0.0671	0.7441	0.0632	0.8181	0.0566
GradientBoosting	0.9417	0.0203	0.8551	0.1058	0.7784	0.0456	0.8124	0.0614
XGBoost	0.9408	0.0166	0.8796	0.0858	0.7387	0.0792	0.7992	0.0589

2.7 MutualInformation2

Table 17: Results for MutualInformation2 on Simple Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.8999	0.0308	0.7000	0.1207	0.6987	0.0671	0.6944	0.0803
RandomForest	0.9235	0.0199	0.8218	0.0954	0.6757	0.0657	0.7391	0.0670
GradientBoosting	0.9199	0.0193	0.7973	0.0943	0.6816	0.0351	0.7333	0.0555
XGBoost	0.9217	0.0248	0.8062	0.1206	0.6986	0.0314	0.7446	0.0628

2.8 ParkingSensores

Table 18: Results for ParkingSensores on Simple Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.8680	0.0172	0.5884	0.0584	0.5905	0.0584	0.5887	0.0553
RandomForest	0.9099	0.0209	0.8019	0.0990	0.5852	0.0586	0.6760	0.0724
GradientBoosting	0.8981	0.0192	0.7192	0.0871	0.6075	0.0634	0.6563	0.0618
XGBoost	0.9035	0.0161	0.7402	0.0769	0.6246	0.0482	0.6752	0.0469

2.9 RFE1

Table 19: Results for RFE1 on Simple Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.9235	0.0207	0.7705	0.0770	0.7495	0.0568	0.7590	0.0626
RandomForest	0.9418	0.0190	0.8773	0.0948	0.7500	0.0609	0.8057	0.0607
GradientBoosting	0.9427	0.0167	0.8843	0.0790	0.7446	0.0525	0.8068	0.0540
XGBoost	0.9418	0.0176	0.8625	0.0757	0.7614	0.0460	0.8080	0.0550

2.10 RFE2

Table 20: Results for RFE2 on Simple Dataset

Model	Accuracy	STD	Precision	STD	Recall	STD	F1-Score	STD
DecisionTree	0.8972	0.0229	0.6849	0.0871	0.6932	0.0855	0.6838	0.0638
RandomForest	0.9281	0.0117	0.8565	0.0937	0.6760	0.0300	0.7519	0.0292
GradientBoosting	0.9217	0.0143	0.8014	0.1036	0.7048	0.0679	0.7431	0.0396
XGBoost	0.9245	0.0128	0.8493	0.1113	0.6652	0.0515	0.7393	0.0319