Evaluation Report

Datasets summary:

Name	Link	Instances	Attributes	Missing values?
Tic Tac Toe	URL	958	10	0
Lymphography	URL	148	19	0
Chess	URL	3196	36	0
Car Evaluation	URL	1728	7	0
Breast Cancer	URL	286	10	9

Performance table:

Dataset	Method	Evaluation metric	Evaluation Value	Fit Runtime (in ms)
Tic Tac Toe	ID3	accuracy	0.9410	17
Tic Tac Toe	MyID3	accuracy	0.9358	1661
Tic Tac Toe	BaggingID3	accuracy	0.9645	171
Tic Tac Toe	MyBaggingID3	accuracy	0.9640	12590
Tic Tac Toe	ID3	precision	0.9495	22
Tic Tac Toe	MyID3	precision	0.9443	1676
Tic Tac Toe	BaggingID3	precision	0.9664	176
Tic Tac Toe	MyBaggingID3	precision	0.9696	12616
Tic Tac Toe	ID3	recall	0.9501	21
Tic Tac Toe	MyID3	recall	0.9458	1681
Tic Tac Toe	BaggingID3	recall	0.9702	177
Tic Tac Toe	MyBaggingID3	recall	0.9696	12766
Tic Tac Toe	ID3	f1_score	0.9512	21
Tic Tac Toe	MyID3	f1_score	0.9470	1683
Tic Tac Toe	BaggingID3	f1_score	0.9709	178
Tic Tac Toe	MyBaggingID3	f1_score	0.9706	12646
Tic Tac Toe	ID3	roc_auc_score	0.9494	22
Tic Tac Toe	MyID3	roc_auc_score	0.9436	1679
Tic Tac Toe	BaggingID3	roc_auc_score	0.9753	175
Tic Tac Toe	MyBaggingID3	roc_auc_score	0.9753	12809
Lymphography	ID3	accuracy	0.8135	11
Lymphography	MyID3	accuracy	0.8245	553
Lymphography	BaggingID3	accuracy	0.8280	128
Lymphography	MyBaggingID3	accuracy	0.8448	3297

Lymphography	ID3	precision	0.7987	14
Lymphography	MyID3	precision	0.8121	563
Lymphography	BaggingID3	precision	0.8293	133
Lymphography	MyBaggingID3	precision	0.8405	3419
Lymphography	ID3	recall	0.7997	16
Lymphography	MyID3	recall	0.8076	573
Lymphography	BaggingID3	recall	0.8027	130
Lymphography	MyBaggingID3	recall	0.8214	3279
Lymphography	ID3	f1_score	0.7967	14
Lymphography	MyID3	f1_score	0.8042	555
Lymphography	BaggingID3	f1_score	0.7976	131
Lymphography	MyBaggingID3	f1_score	0.8211	3326
Lymphography	ID3	roc_auc_score	0.8003	13
Lymphography	MyID3	roc_auc_score	0.8080	554
Lymphography	BaggingID3	roc_auc_score	0.8198	129
Lymphography	MyBaggingID3	roc_auc_score	0.8379	3226
Chess	ID3	accuracy	0.9812	60
Chess	MyID3	accuracy	0.9826	10258
Chess	BaggingID3	accuracy	0.9876	370
Chess	MyBaggingID3	accuracy	0.9834	74487
Chess	ID3	precision	0.9775	65
Chess	MyID3	precision	0.9771	10377
Chess	BaggingID3	precision	0.9859	382
Chess	MyBaggingID3	precision	0.9891	77029
Chess	ID3	recall	0.9819	66
Chess	MyID3	recall	0.9833	10224
Chess	BaggingID3	recall	0.9880	380
Chess	MyBaggingID3	recall	0.9827	73844
Chess	ID3	f1_score	0.9820	65
Chess	MyID3	f1_score	0.9834	10210
Chess	BaggingID3	f1_score	0.9881	378
Chess	MyBaggingID3	f1_score	0.9830	75185
Chess	ID3	roc_auc_score	0.9850	67
Chess	MyID3	roc_auc_score	0.9831	10279
Chess	BaggingID3	roc_auc_score	0.9901	381
Chess	MyBaggingID3	roc_auc_score	0.9861	72265
Car Evaluation	ID3	accuracy	0.9679	17
Car Evaluation	MyID3	accuracy	0.9676	1633
Car Evaluation	BaggingID3	accuracy	0.9766	169
Car Evaluation	MyBaggingID3	accuracy	0.9705	13327
Car Evaluation	ID3	precision	0.9772	23
Car Evaluation	MyID3	precision	0.9797	1627
Car Evaluation	BaggingID3	precision	0.9840	182
Car Evaluation	MyBaggingID3	precision	0.9852	12620
Car Evaluation	ID3	recall	0.9740	22
Car Evaluation	MyID3	recall	0.9737	1644
Car Evaluation	BaggingID3	recall	0.9810	173

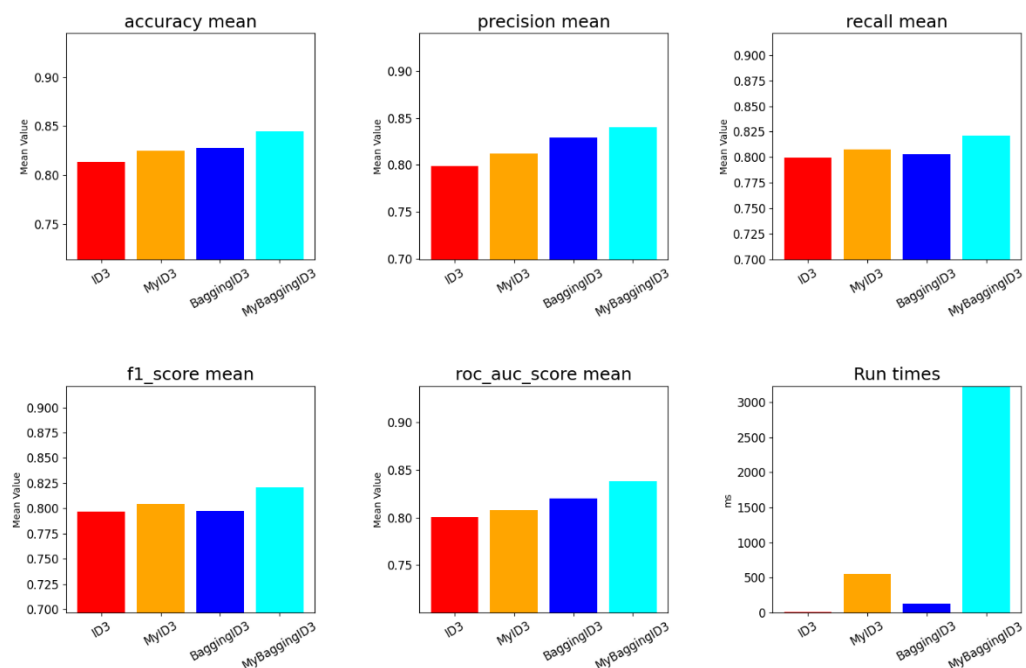|				
Car Evaluation	MyBaggingID3	recall	0.9762	13679
Car Evaluation	ID3	f1_score	0.9747	22
Car Evaluation	MyID3	f1_score	0.9744	1644
Car Evaluation	BaggingID3	f1_score	0.9815	174
Car Evaluation	MyBaggingID3	f1_score	0.9769	12734
Car Evaluation	ID3	roc_auc_score	0.9769	22
Car Evaluation	MyID3	roc_auc_score	0.9739	1646
Car Evaluation	BaggingID3	roc_auc_score	0.9844	173
Car Evaluation	MyBaggingID3	roc_auc_score	0.9807	13491
Breast Cancer	ID3	accuracy	0.6969	13
Breast Cancer	MyID3	accuracy	0.7058	1626
Breast Cancer	BaggingID3	accuracy	0.7238	139
Breast Cancer	MyBaggingID3	accuracy	0.7383	7830
Breast Cancer	ID3	precision	0.5918	16
Breast Cancer	MyID3	precision	0.6014	1626
Breast Cancer	BaggingID3	precision	0.6424	141
Breast Cancer	MyBaggingID3	precision	0.6672	7953
Breast Cancer	ID3	recall	0.5220	17
Breast Cancer	MyID3	recall	0.5256	1637
Breast Cancer	BaggingID3	recall	0.5370	147
Breast Cancer	MyBaggingID3	recall	0.5353	8047
Breast Cancer	ID3	f1_score	0.4957	16
Breast Cancer	MyID3	f1_score	0.5002	1642
Breast Cancer	BaggingID3	f1_score	0.5030	144
Breast Cancer	MyBaggingID3	f1_score	0.4899	7937
Breast Cancer	ID3	roc_auc_score	0.5210	16
Breast Cancer	MyID3	roc_auc_score	0.5219	1628
Breast Cancer	BaggingID3	roc_auc_score	0.5419	143
Breast Cancer	MyBaggingID3	roc_auc_score	0.5287	8105

The averages per method are:

	ID3	MyID3	Bagging	MyBagging
Accuracy	0.8801	0.8833	0.8961	0.9002
Precision	0.8401	0.8418	0.8482	0.8483
Recall	0.8589	0.8629	0.8816	0.8903
F1 score	0.8455	0.8472	0.8558	0.8570
ROC AUC	0.8465	0.8461	0.8623	0.8617
Runtime	27	3155	200	22340

Comparison histograms

Tic Tac Toe

accuracy mean

precision mean

recall mean

f1_score mean

roc_auc_score mean

Run times

Lymphography

accuracy mean

precision mean

recall mean
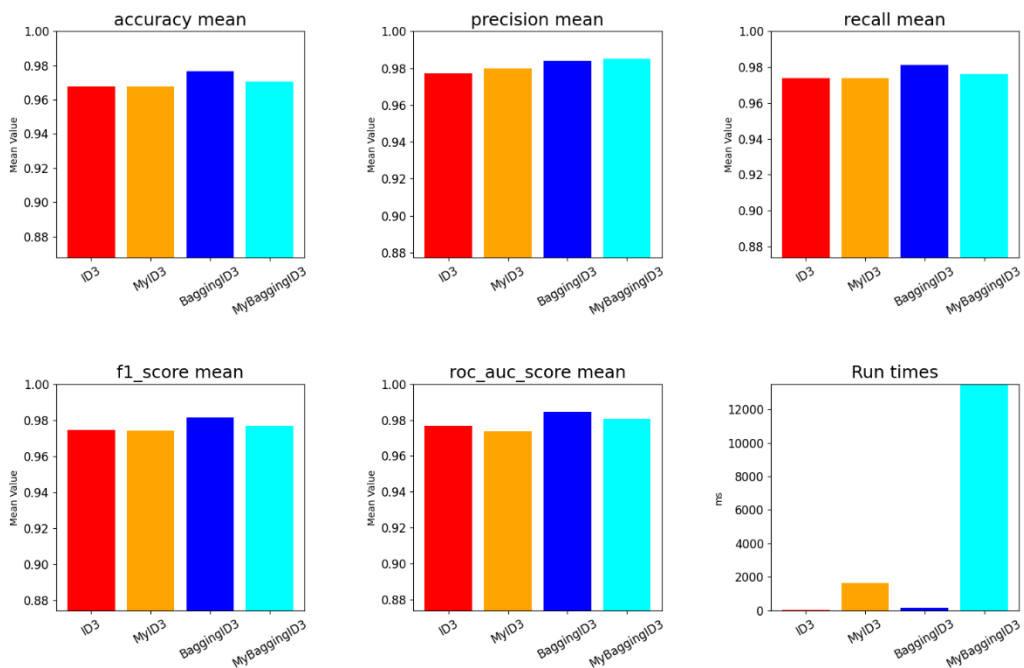
f1_score mean

roc_auc_score mean
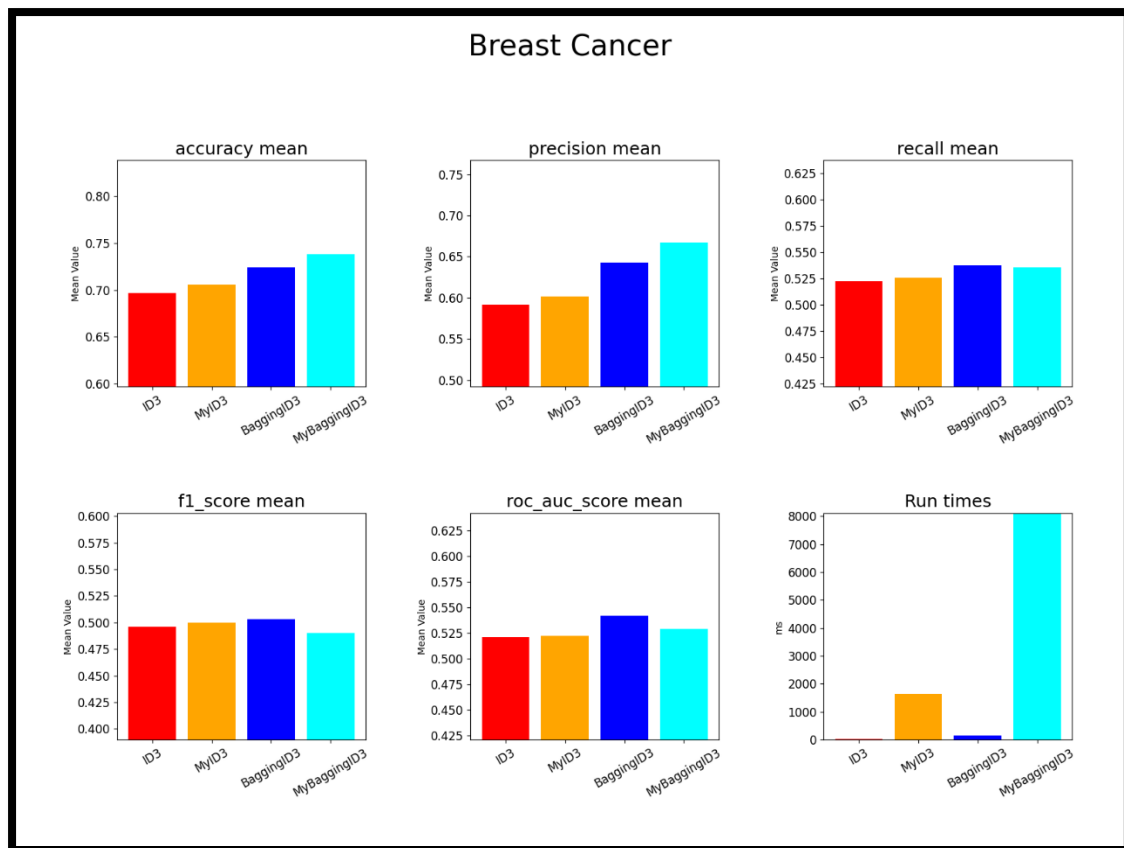
Run times
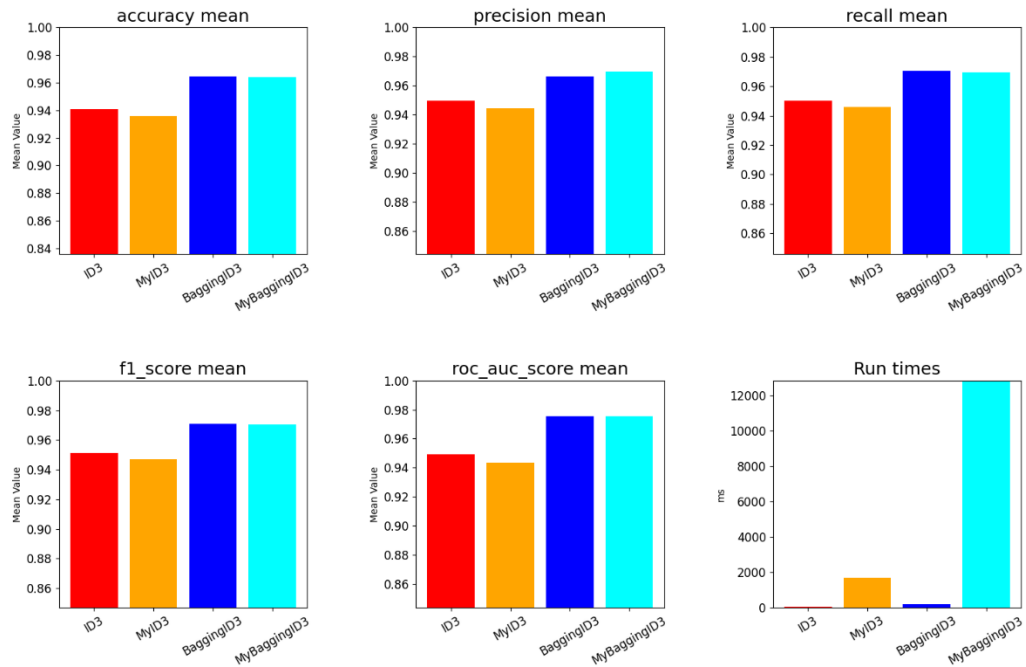
Chess kr-vs-kp

Car Evaluation

Breast Cancer
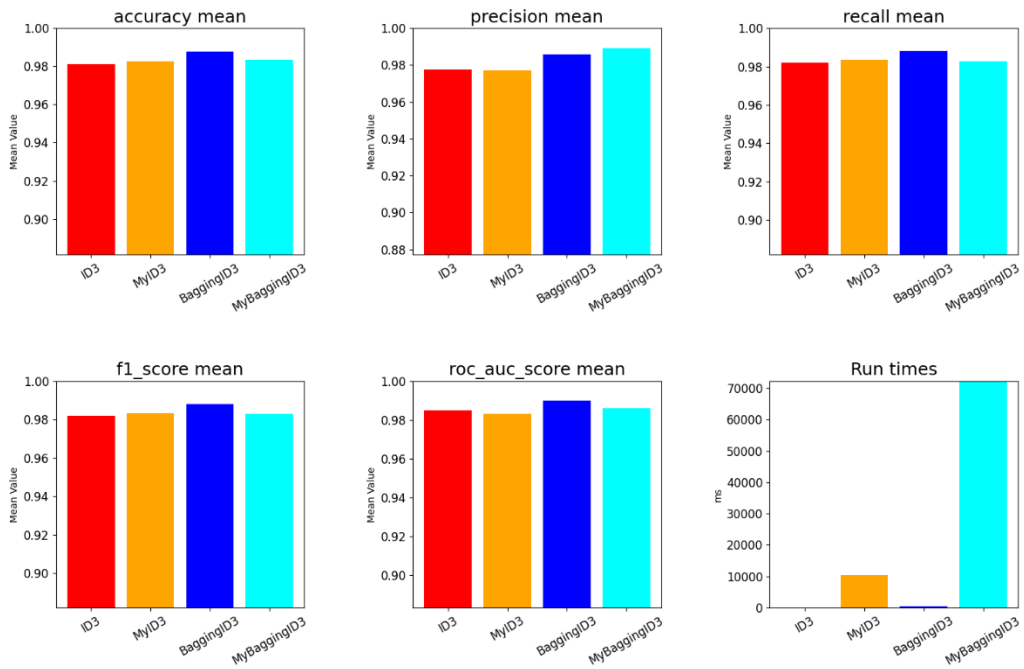
Summary:

- Our implementation of ID3 provided better results across the following metrics: accuracy, precision, recall, F1 score.

- Our implementation of BaggingID3 also provided better results across the following metrics: accuracy, precision, recall, F1 score.

- The best performing method overall was the sklearn implementation of BaggingID3, since its runtime was significantly lower than our implementation.

- The runtime of our methods compared to the ones from the sklearn package is around 100 times longer, the performance of our implementations across the different metrics is only slightly better (around 0.3% improvement on average).