Sampling Percent	Algorithm	Benchmarks								Avg Wins
		breast_cancer	churn	JS_Vuln	Ambari_Vuln	Defect_Eclipse_JDT	Defect_Eclipse_PDE	Moodle_Vuln	Defect_Mylyn	
20%	random_oversampling	0.40	0.73	0.50	0.24	0.50	0.29	0.00	0.33	0.50
	smote	0.40	0.72	0.51	0.22	0.54	0.36	0.00	0.35	1.13
	svm_smote	0.40	0.71	0.51	0.22	0.49	0.29	0.00	0.35	0.50
	gaussian_copula	0.34	0.67	0.50	0.14	0.47	0.00	0.00	0.30	0.25
	RRP	0.43	0.69	0.49	0.36	0.51	0.31	0.00	0.36	0.38
	random_pruning	0.41	0.71	0.51	0.17	0.49	0.35	0.00	0.34	0.50
	intelligent_pruning	0.41	0.71	0.51	0.17	0.50	0.34	0.09	0.33	1.50
40%	random_oversampling	0.38	0.73	0.52	0.29	0.48	0.27	0.00	0.32	0.50
	smote	0.39	0.72	0.51	0.15	0.53	0.30	0.00	0.36	1.00
	svm_smote	0.38	0.72	0.50	0.15	0.51	0.32	0.00	0.36	0.50
	gaussian_copula	0.40	0.67	0.51	0.18	0.48	0.29	0.00	0.30	0.25
	RRP	0.46	0.69	0.49	0.20	0.52	0.34	0.00	0.33	0.50
	random_pruning	0.39	0.69	0.48	0.24	0.48	0.30	0.00	0.33	0.25
	intelligent_pruning	0.44	0.70	0.49	0.29	0.53	0.32	0.13	0.36	1.50
60%	random_oversampling	0.40	0.73	0.50	0.24	0.51	0.28	0.00	0.32	0.50
	smote	0.40	0.72	0.51	0.22	0.54	0.31	0.00	0.35	1.13
	svm_smote	0.40	0.72	0.51	0.22	0.52	0.33	0.00	0.35	0.50
	gaussian_copula	0.34	0.67	0.50	0.15	0.47	0.29	0.00	0.30	0.25
	RRP	0.43	0.69	0.49	0.36	0.52	0.30	0.00	0.36	0.38
	random_pruning	0.41	0.67	0.51	0.25	0.53	0.35	0.00	0.34	0.50
	intelligent_pruning	0.41	0.65	0.51	0.27	0.52	0.34	0.09	0.33	1.50
80%	random_oversampling	0.41	0.73	0.50	0.24	0.49	0.26	0.00	0.28	0.50
	smote	0.46	0.72	0.52	0.10	0.51	0.37	0.00	0.36	1.00
	svm_smote	0.46	0.72	0.51	0.18	0.51	0.30	0.00	0.36	0.50
	gaussian_copula	0.40	0.67	0.49	0.15	0.45	0.29	0.00	0.32	0.13
	RRP	0.42	0.68	0.49	0.17	0.54	0.34	0.00	0.32	0.38
	random_pruning	0.43	0.62	0.46	0.18	0.52	0.32	0.00	0.37	0.38
	intelligent_pruning	0.44	0.65	0.49	0.22	0.52	0.35	0.10	0.35	1.88
100%	random_oversampling	0.36	0.72	0.53	0.24	0.15	0.30	0.00	0.32	0.5
	smote	0.38	0.71	0.51	0.11	0.15	0.33	0.00	0.36	0.87
	svm_smote	0.38	0.70	0.50	0.17	0.15	0.32	0.00	0.36	0.37
	gaussian_copula	0.39	0.67	0.50	0.15	0.47	0.29	0.00	0.30	0.25
	RRP	0.41	0.67	0.49	0.20	0.50	0.32	0.00	0.35	0.5
	random_pruning	0.40	0.54	0.42	0.12	0.54	0.32	0.00	0.31	0.37
	intelligent_pruning	0.47	0.65	0.46	0.18	0.50	0.35	0.12	0.34	1.62

Table 1: **Model comparison scores for different sampling ratios for DT. All the tasks are evaluated by F1-Score using different sampling algorithms for various sampling ratios. Higher values are better. The dark grey cells mark the algorithms in the Rank 0 of the Skott Knott plots. Avg Wins refers to the average number of times a sampling algorithm for a fixed sampling ratio came in Rank 0 across all the 8 datasets. Due to space constraints Recursive Random Projection is simplified to RRP. In the case of random & intelligent pruning we undersample the dataset, in no-sampling we don't do any sampling and by rest of the techniques we oversample.**

v

Sampling Percent	Algorithm	Benchmarks								Avg Wins
		breast_cancer	churn	JS_Vuln	Ambari_Vuln	Defect_Eclipse_JDT	Defect_Eclipse_PDE	Moodle_Vuln	Defect_Mylyn	
20%	random_oversampling	0.45	0.42	0.39	0.00	0.55	0.38	0.13	0.30	0.12
	smote	0.45	0.44	0.39	0.00	0.56	0.38	0.15	0.31	0.50
	svm_smote	0.45	0.46	0.39	0.00	0.56	0.36	0.15	0.30	0.375
	gaussian_copula	0.44	0.37	0.29	0.00	0.56	0.31	0.00	0.22	0.00
	RRP	0.45	0.41	0.25	0.1	0.57	0.35	0.11	0.29	0.37
	random_pruning	0.45	0.34	0.32	0.00	0.55	0.31	0.00	0.26	0.00
	intelligent_pruning	0.44	0.42	0.42	0.00	0.59	0.38	0.12	0.27	0.625
40%	random_oversampling	0.48	0.46	0.46	0.00	0.57	0.38	0.06	0.34	0.25
	smote	0.47	0.45	0.47	0.06	0.57	0.40	0.06	0.36	0.75
	svm_smote	0.47	0.48	0.45	0.00	0.56	0.39	0.11	0.31	0.37
	gaussian_copula	0.47	0.41	0.31	0.00	0.54	0.31	0.00	0.24	0.125
	RRP	0.48	0.46	0.39	0.08	0.58	0.38	0.07	0.35	0.375
	random_pruning	0.45	0.39	0.35	0.00	0.55	0.30	0.00	0.29	0.0
	intelligent_pruning	0.45	0.42	0.43	0.12	0.61	0.38	0.10	0.31	0.625
60%	random_oversampling	0.47	0.49	0.48	0.00	0.56	0.40	0.07	0.37	0.62
	smote	0.44	0.48	0.50	0.04	0.56	0.40	0.07	0.36	0.62
	svm_smote	0.44	0.48	0.49	0.00	0.55	0.40	0.08	0.35	0.25
	gaussian_copula	0.47	0.42	0.31	0.00	0.53	0.30	0.00	0.24	0.12
	RRP	0.50	0.49	0.45	0.08	0.57	0.41	0.04	0.37	0.75
	random_pruning	0.49	0.44	0.40	0.00	0.59	0.35	0.00	0.34	0.25
	intelligent_pruning	0.47	0.44	0.45	0.12	0.60	0.38	0.14	0.31	1.12
80%	random_oversampling	0.51	0.49	0.46	0.00	0.55	0.40	0.07	0.37	0.38
	smote	0.44	0.49	0.48	0.04	0.56	0.41	0.08	0.37	0.75
	svm_smote	0.42	0.49	0.48	0.00	0.56	0.40	0.08	0.37	0.38
	gaussian_copula	0.49	0.41	0.34	0.00	0.53	0.29	0.00	0.23	0.13
	RRP	0.50	0.49	0.44	0.04	0.55	0.41	0.07	0.38	0.50
	random_pruning	0.49	0.47	0.46	0.08	0.55	0.36	0.00	0.33	0.25
	intelligent_pruning	0.51	0.46	0.48	0.10	0.59	0.41	0.16	0.33	2.13
100%	random_oversampling	0.50	0.48	0.44	0.00	0.55	0.41	0.07	0.36	0.38
	smote	0.43	0.48	0.43	0.04	0.53	0.41	0.07	0.37	0.50
	svm_smote	0.42	0.47	0.44	0.00	0.55	0.40	0.06	0.36	0.25
	gaussian_copula	0.49	0.40	0.34	0.00	0.53	0.31	0.00	0.22	0.13
	RRP	0.47	0.48	0.41	0.08	0.55	0.41	0.08	0.37	0.50
	random_pruning	0.47	0.48	0.43	0.08	0.53	0.41	0.06	0.34	0.38
	intelligent_pruning	0.52	0.45	0.48	0.06	0.57	0.41	0.15	0.36	1.25

Table 2: **Model comparison scores for different sampling ratios for LR. All the tasks are evaluated by F1-Score using different sampling algorithms for various sampling ratios. Higher values are better. The dark grey cells mark the algorithms in the Rank 0 of the Skott Knott plots. Avg Wins refers to the average number of times a sampling algorithm for a fixed sampling ratio came in Rank 0 across all the 8 datasets. Due to space constraints Recursive Random Projection is simplified to RRP. In the case of random & intelligent pruning we undersample the dataset, in no-sampling we don't do any sampling and by rest of the techniques we oversample.**