

Features	Pd		Precision		F		SUM	
	naive	tuned	naive	tuned	naive	tuned	naive	tuned
noc								
ca								
max_cc			1		1		2	
ce		1	1		2		3	
moa			1		2		4	
cbo			1		4		5	
avg_cc			3		2		5	
lcom			3		3		6	
npm		1	5		4		10	
cbm	4		6	3	4	3	14	6
amc	4	1	4	3	4	5	12	9
rfc	4	1	4	5	4	9	12	15
ic	8		7	4	9	4	24	8
wmc	5	2	5	6	5	11	15	19
dit	8		8	6	8	6	24	12
lcom3	9	1	9	7	9	7	27	15
loc	9	2	8	6	9	12	26	20
cam	9	1	9	8	9	10	27	19
dam	14	3	14	9	14	11	42	23
mfa	16	3	16	11	16	13	48	27

Figure 9: Counts of features selected by different goals. For each goal, the numbers in right and left columns represent the counts of features selected for all the data sets with and without tuning processes.

Data set	WHERE		CART		Random Forests	
	Naive	Tuned	Naive	Tuned	Naive	Tuned
antV0	53	★100	38	★100	97	★100
antV1	7	★100	39	96	95	★100
antV2	0	★100	37	★100	92	98
camelV0	0	★100	6	★100	66	97
camelV1	80	★100	46	★100	94	★100
ivyV0	93	★100	88	★100	98	98
jeditV0	89	★100	66	99	96	★100
jeditV1	75	★100	50	★100	★100	★100
jeditV2	45	★100	36	★100	★100	★100
log4jV0	46	★100	31	★100	79	94
luceneV0	81	★100	48	★100	★100	★100
poiV0	89	96	79	★100	99	★100
poiV1	2	★100	12	★100	89	★100
synapseV0	0	★100	28	★100	90	★100
velocityV0	★100	★100	86	★100	99	★100
xercesV0	64	★100	46	★100	81	★100
xercesV1	15	★100	11	★100	68	88

Figure 10: pd values in tune and naive runs. All data available from <http://openscience.us/repo/effort>. KEY: percentile ranges: 80th to 100th= ■■■; 60th to 80th= ■■; 40th to 60th= ■; 20th to 40th= □; an absent bar shows 0th to 20th. Percentiles computed separately for each row.

Data set	WHERE		CART		Random Forests	
	Naive	Tuned	Naive	Tuned	Naive	Tuned
antV0	30	★89	27	★89	39	★89
antV1	32	★74	41	★74	43	0
antV2	★78	★78	52	68	66	67
camelV0	★83	★83	26	33	34	45
camelV1	22	★28	23	24	★28	★28
ivyV0	16	★23	18	21	21	20
jeditV0	35	★75	49	56	50	48
jeditV1	24	★87	28	86	36	39
jeditV2	2	★98	3	18	5	5
log4jV0	94	★100	97	★100	★100	★100
luceneV0	61	74	67	70	72	★77
poiV0	70	68	77	72	★79	76
poiV1	★100	90	73	89	81	★100
synapseV0	66	66	71	★100	59	80
velocityV0	34	39	34	40	40	★41
xercesV0	13	★85	14	13	15	13
xercesV1	★56	26	50	26	41	26

Figure 11: precision values in tune and naive runs. All data available from <http://openscience.us/repo/effort>. KEY: percentile ranges: 80th to 100th= ■■■; 60th to 80th= ■■; 40th to 60th= ■; 20th to 40th= □; an absent bar shows 0th to 20th. Percentiles computed separately for each row.

Data set	WHERE		CART		Random Forests	
	Naive	Tuned	Naive	Tuned	Naive	Tuned
antV0	★39	25	32	31	★39	12
antV1	11	6	40	★45	39	44
antV2	0	★87	44	1	50	51
camelV0	0	28	9	28	★34	30
camelV1	★34	★34	31	32	33	31
ivyV0	27	34	30	★38	35	33
jeditV0	50	56	56	54	★61	59
jeditV1	37	33	36	★49	45	47
jeditV2	4	8	5	★11	9	9
log4jV0	★62	56	47	59	53	43
luceneV0	70	★75	56	★75	73	★75
poiV0	★78	64	74	68	73	69
poiV1	5	★78	21	★78	76	★78
synapseV0	0	2	40	★56	51	55
velocityV0	★51	★51	49	★51	★51	★51
xercesV0	22	20	21	★27	23	20
xercesV1	25	39	18	53	68	★73

Figure 12: f values in tune and naive runs. All data available from <http://openscience.us/repo/effort>. KEY: percentile ranges: 80th to 100th= ■■■; 60th to 80th= ■■; 40th to 60th= ■; 20th to 40th= □; an absent bar shows 0th to 20th. Percentiles computed separately for each row.

Learner Name	Parameters	Default	antV0	antV1	antV2	camelV0	camelV1	ivyV0	jeditV0	jeditV1	jeditV2	log4jV0	luceneV0	poiV0	poiV1	synapseV0	velocityV0	xercesV0	xercesV1
Where based Learner	threshold	0.5	0.17	0.14	0.01	0.03	0.4	0.03	0.44	0.14	0.29	0.09	0.56	0.56	0.01	0.01	0.85	0.31	0.07
	infoPrune	0.33	0.13	0.39	0.01	0.17	0.04	0.44	0.6	0.62	0.96	0.91	0.44	0.4	0.1	0.07	0.89	0.47	0.34
	min_sample_split	4	5	1	9	1	5	8	6	2	8	8	1	2	8	5	9	2	1
	min.Size	0.5	0.49	0.15	0.49	0.27	0.15	0.99	0.13	0.24	0.27	0.07	0.1	0.98	0.56	0.58	0.27	0.18	0.21
	wriggle	0.2	0.65	0.72	0.86	0.53	0.52	0.42	0.93	0.53	0.35	0.08	0.37	0.47	0.01	0.74	0.68	0.02	0.99
	depthMin	2	4	5	6	2	4	3	3	1	4	5	4	2	1	4	1	4	2
	depthMax	10	16	20	3	18	10	13	10	9	14	10	8	4	12	4	13	10	13
	wherePrune	False	True	False	False	False	True	False	True	False	True	True	True	False	True	False	True	False	False
	treePrune	True	True	False	False	True	True	False	True	False	True	True	False	True	False	False	True	True	True
	treePrune	True	True	False	False	True	True	False	True	False	True	True	False	True	False	False	True	True	True
CART	threshold	0.5	0.01	0.01	0.01	0.01	0.01	0.17	0.01	0.01	0.01	0.18	0.52	0.02	0.01	0.01	0.01	0.06	0.01
	max_feature	None	0.01	0.01	0.53	1	0.27	0.12	0.01	0.27	0.63	0.11	0.07	0.13	0.01	0.71	0.3	0.14	0.02
	min_samples_split	2	19	7	10	20	18	11	2	2	10	4	6	5	20	6	8	15	2
	min_samples_leaf	1	3	12	1	1	3	10	17	16	8	10	13	17	13	3	16	19	12
	threshold	0.5	0.01	0.01	0.01	0.01	0.16	0.23	0.04	0.04	0.05	0.03	0.01	0.07	0.01	0.01	0.26	0.11	0.01
Random Forests	max_feature	None	0.14	0.78	0.96	0.01	0.14	1.0	0.41	0.09	0.91	0.39	0.01	0.82	0.01	0.01	0.29	0.46	0.32
	max_leaf_nodes	None	49	23	27	41	22	16	16	14	30	35	17	44	47	48	20	44	44
	min_samples_split	2	9	9	10	20	3	17	2	5	16	5	19	12	4	17	9	14	9
	min_samples_leaf	1	16	15	9	20	10	8	19	7	17	18	5	15	18	9	11	17	14
	n_estimators	100	140	102	55	108	140	70	107	147	65	145	96	142	133	82	88	119	104

Figure 1: Parameters tuned on different models over the objective of pd

Learner Name	Parameters	Default	antV0	antV1	antV2	camelV0	camelV1	ivyV0	jeditV0	jeditV1	jeditV2	log4jV0	luceneV0	poiV0	poiV1	synapseV0	velocityV0	xercesV0	xercesV1
Where based Learner	threshold	0.5	0.53	0.57	0.14	0.67	0.9	1	1	0.72	0.91	0.96	0.67	1.0	0.41	0.24	1	0.47	0.74
	infoPrune	0.33	0.68	0.45	0.53	0.26	0.81	0.42	0.39	0.21	0.05	0.57	0.35	0.53	0.68	0.52	0.56	0.5	0.75
	min_sample_split	4	3	9	9	8	6	5	2	8	6	1	3	7	1	4	6	9	7
	min.Size	0.5	0.07	0.26	0.39	0.73	0.29	0.49	0.19	0.44	0.36	1.0	0.93	0.45	0.52	0.22	0.91	0.19	0.87
	wriggle	0.2	0.91	0.83	0.31	0.18	0.97	0.74	0.98	0.59	0.61	0.34	0.92	0.71	1	0.35	0.25	0.01	0.17
	depthMin	2	4	5	1	5	3	1	4	3	3	3	1	3	1	2	5	3	2
	depthMax	10	10	17	6	6	12	19	14	15	12	12	18	19	11	8	10	11	18
	wherePrune	False	True	True	True	False	True	True	True	True	False	False	True	True	True	True	False	True	True
	treePrune	True	True	True	False	False	False	True	True	False	False	False	False	False	True	False	False	True	False
	treePrune	True	True	True	False	False	False	True	True	False	False	False	False	False	True	False	False	True	False
CART	threshold	0.5	0.83	1	0.99	0.85	0.9	0.93	1	0.98	0.88	0.9	0.7	0.63	0.63	0.01	0.02	0.19	0.28
	max_feature	None	0.01	0.35	0.35	0.73	0.63	0.76	0.61	0.13	0.01	0.83	0.01	0.02	0.19	0.28	0.97	0.1	0.17
	min_samples_split	2	18	11	6	18	20	19	3	6	14	12	8	12	17	2	11	5	16
	min_samples_leaf	1	3	19	6	1	4	2	14	1	2	6	2	19	4	6	15	1	17
	threshold	0.5	0.86	0.61	0.71	0.17	1	1	0.87	1	1	1	0.97	1	0.41	0.33	1	0.83	0.94
Random Forests	max_feature	None	0.05	0.34	0.31	0.45	0.54	0.94	0.93	0.81	0.01	1	0.69	0.42	1	0.01	0.01	0.75	1
	max_leaf_nodes	None	16	11	38	47	18	38	27	33	16	24	41	50	15	15	30	47	28
	min_samples_split	2	3	3	4	19	17	19	18	5	13	19	1	5	2	6	7	11	11
	min_samples_leaf	1	7	11	10	16	8	14	4	2	7	14	12	7	10	12	7	8	4
	n_estimators	100	85	111	99	121	128	140	141	63	100	69	99	109	123	88	125	107	131

Figure 2: Parameters tuned on different models over the objective of prec

Learner Name	Parameters	Default	antV0	antV1	antV2	camelV0	camelV1	ivyV0	jeditV0	jeditV1	jeditV2	log4jV0	luceneV0	poiV0	poiV1	synapseV0	velocityV0	xercesV0	xercesV1
Where based Learner	threshold	0.5	0.02	0.72	0.97	0.01	0.87	1	0.59	0.49	0.77	0.38	0.25	0.94	0.01	0.81	0.53	0.83	0.99
	infoPrune	0.33	0.13	0.64	0.57	0.33	0.92	0.37	0.43	0.53	0.68	0.56	0.55	0.74	0.04	0.34	0.62	1	0.36
	min_sample_split	4	5	1	1	5	3	8	9	2	4	6	2	2	2	2	2	9	7
	min.Size	0.5	0.59	0.56	0.29	0.03	0.64	0.5	1.0	0.45	1	0.79	0.22	0.97	1	0.6	0.68	0.47	0.35
	wriggle	0.2	0.51	0.88	0.92	0.67	0.81	0.82	0.45	0.93	0.39	0.92	0.16	0.48	0.22	0.98	0.7	0.57	0.02
	depthMin	2	4	2	5	3	2	1	5	2	2	1	3	3	3	1	4	2	5
	depthMax	10	16	12	1	11	19	12	20	7	8	17	19	17	12	9	9	13	3
	wherePrune	False	True	True	False	False	False	True	False	False	True	False	False	True	False	False	False	False	False
	treePrune	True	True	False	False	False	True	False	False	False	True	True	False	False	False	False	True	False	False
	treePrune	True	True	False	False	False	True	False	False	False	True	True	False	False	False	False	True	False	False
CART	threshold	0.5	0.06	0.63	0.99	0.01	1	0.98	0.47	0.74	0.99	0.36	0.6	0.95	0.01	0.01	0.01	0.43	0.01
	max_feature	None	0.3	0.81	0.01	1	0.98	0.26	0.36	0.76	0.29	0.5	0.12	0.01	0.01	0.01	0.16	0.96	0.01
	min_samples_split	2	13	10	15	14	13	7	7	4	5	5	15	20	8	20	6	10	2
	min_samples_leaf	1	9	17	1	9	1	20	1	20	17	4	9	3	12	10	16	4	8
	threshold	0.5	0.92	0.49	0.24	0.01	1	1	1	1	1	0.7	0.21	1	0.01	0.17	0.4	0.81	0.01
Random Forests	max_feature	None	0.02	0.55	0.01	0.01	0.81	0.52	0.01	0.76	0.01	0.01	0.67	0.3	0.01	0.06	0.36	0.34	0.93
	max_leaf_nodes	None	13	10	34	10	15	25	34	24	40	20	33	14	27	40	47	44	41
	min_samples_split	2	8	1	15	15	10	11	10	5	1	9	6	7	13	13	12	15	15
	min_samples_leaf	1	2	14	17	19	5	13	5	2	3	11	17	2	5	9	7	18	14
	n_estimators	100	132	50	81	150	63	59	85	141	96	106	71	95	121	135	112	104	92

Figure 3: Parameters tuned on different models over the objective of F

Dataset	antV0	antV1	antV2	camelV0	camelV1	ivyV0	jeditV0	jeditV1	jeditV2
training	20/125	40/178	32/293	13/339	216/608	63/111	90/272	75/306	79/312
tunning	40/178	32/293	92/351	216/608	145/872	16/241	75/306	79/312	48/367
testing	32/293	92/351	166/745	145/872	188/965	40/352	79/312	48/367	11/492

Figure 4: The percentage of defective instances in each experimental data set. For each experiment, training, tuning and testing data are composed of single chronological data file

Dataset	log4jV0	luceneV0	poiV0	poiV1	synapseV0	velocityV0	xercesV0	xercesV1
training	34/135	91/195	141/237	37/314	16/157	147/196	77/162	71/440
tunning	37/109	144/247	37/314	248/385	60/222	142/214	71/440	69/453
testing	189/205	203/340	248/385	281/442	86/256	78/229	69/453	437/588

Figure 5: The percentage of defective instances in each experimental data set. For each experiment, training, tuning and testing data are composed of single chronological data file

Features	antV0	antV1	antV2	camelV0	camelV1	ivyV0	jeditV0	jeditV1	jeditV2	log4jV0	luceneV0	poiV0	poiV1	synapseV0	velocityV0	xercesV0	xercesV1
npm		*		*	*					*		o	o	*		*	*
loc												o	o				
amc												*	o	*		*	*
max_cc																	
lcom																	
dam	*	*	*	*		*	*	*	*		*	*	*	*	*	*	
ca																	
cbo																	
ce																	
noc																	
rfc				*	*		o					*					*
dit		*		*	*	*	*	*	*	*					*		
mfa	*	o	*	*	*	*	*	*	*	*	*	*	o	*	*	*	*
cam			*	*	*	*	o	*	*		*			*		*	*
avg_cc																	
wmc	*	*			*		o					*	o				*
moa							o										
cbm								*	*					*			
ic	*		*			*		*	*	*	*				*		
lcom3			*	*		*	*	*	*	*	*		o		*		

Figure 6: Feature selection for different datasets with and without the tuning process over the objective of pd.

Features	antV0	antV1	antV2	camelV0	camelV1	ivyV0	jeditV0	jeditV1	jeditV2	log4jV0	luceneV0	poiV0	poiV1	synapseV0	velocityV0	xercesV0	xercesV1
npm		*		*	o	*	o			*	o	o	o	*		*	o
loc												o	o			*	o
amc												*	o	*		*	o
max_cc																	
lcom															o		
dam	*	*	*	o	*		o	*	*	o	*	*	o	*	*	*	o
ca											*	o	*				
cbo																	o
ce							o										
noc																	
rfc				o	*	o	*	*	*		o	*	o		o		*
dit		*		o	*	o	*	*	o	*			o		*		o
mfa	*	*	*	o	*	o	*	*	o	*	o	*	o	*	*	*	o
cam			*	o	*	o	*	*	o	*	o	*	o	*	o	*	o
avg_cc				o					*		o		o				o
wmc	*	*			*						o	*	o	o	o		o
moa															o		o
cbm						*	*	*	*	*		o	*	o			o
ic	*		*			o	*	*	o	*	*	o			*	o	
lcom3			*	*	o	*	o	*	*	*	o	o	*	o	*	o	

Figure 7: Feature selection for different datasets with and without the tuning process over the objective of precision.

Features	antV0	antV1	antV2	camelV0	camelV1	ivyV0	jeditV0	jeditV1	jeditV2	log4jV0	luceneV0	poiV0	poiV1	synapseV0	velocityV0	xercesV0	xercesV1
npm		o	o		*	o		o	o	*	o	*	o	*	o	*	o
loc		*	o	*	*	o		o	o			*	o	*	o	*	o
amc		o	o									*	o	*	o	*	o
max_cc		o	o												o		
lcom		o	o									o					
dam	*	*	o	*		o	*	o	*	o	*	*	*	*	o	*	o
ca											*	*	*	*	o	*	o
cbo						o									o	o	o
ce												o					
noc																	
rfc	o	o	o	*	*	o	*	o	*	o	*	*		o	o	*	o
dit		*	o	*	*	o	*	o	*	o	*	*	*		*	o	o
mfa	*	o	*	o	*	o	*	o	*	o	*	*	*	*	*	*	o
cam			*	o	*	o	*	o	*	o	*	*	*	*	o	*	o
avg_cc				o					o			*	o		o	o	o
wmc	*	*	o		*	o		o	o	o		*	o		o	o	o
moa									o								
cbm					o		*	*	*	*	*		*		o		
ic	*		*		o	*	*	*	*	*	*	o	*		*	o	
lcom3		o	*	*	o	*	*	*	*	*	*	o	*		*	o	

Figure 8: Feature selection for different datasets with and without the tuning process over the objective of F measure.

Data set	WHERE		CART		Random Forests	
	Naive	Tuned	Naive	Tuned	Naive	Tuned
antV0	65	68	53	★75	69	72
antV1	12	★70	52	56	62	69
antV2	0	★69	53	64	68	67
camelV0	0	51	10	56	60	★61
camelV1	45	55	53	★59	56	57
ivy	53	71	63	★74	70	71
jeditV0	59	69	71	72	★73	★73
jeditV1	70	69	62	73	★75	★75
jeditV2	49	56	48	61	★78	67
log4j	53	56	46	★59	56	56
lucene	37	65	55	56	60	★66
poiV0	45	60	66	62	★68	56
poiV1	6	★62	22	60	56	57
synapse	0	61	43	64	63	★67
velocity	3	★56	26	51	51	55
xercesV0	37	46	47	★49	★49	48
xercesV1	24	★68	19	32	32	17

Figure 13: g values in tune and naive runs. All data available from <http://openscience.us/repo/effort>. KEY: percentile ranges: 80th to 100th= ■■■; 60th to 80th= ■■; 40th to 60th= ■; 20th to 40th= □; an absent bar shows 0th to 20th. Percentiles computed separately for each row.

Data set	WHERE		CART		Random Forests	
	Naive	Tuned	Naive	Tuned	Naive	Tuned
antV0	30	★89	27	★89	39	★89
antV1	32	★74	41	★74	43	0
antV2	★78	★78	52	68	66	67
camelV0	★83	★83	26	33	34	45
camelV1	22	★28	23	24	★28	★28
ivyV0	16	★23	18	21	21	20
jeditV0	35	★75	49	56	50	48
jeditV1	24	★87	28	86	36	39
jeditV2	2	★98	3	18	5	5
log4jV0	94	★100	97	★100	★100	★100
luceneV0	61	74	67	70	72	★77
poiV0	70	68	77	72	★79	76
poiV1	★100	90	73	89	81	★100
synapseV0	66	66	71	★100	59	80
velocityV0	34	39	34	40	40	★41
xercesV0	13	★85	14	13	15	13
xercesV1	★56	26	50	26	41	26
antV0	★89	0	15	0	21	0
antV1	0	0	54	0	★68	0
antV2	53	0	43	0	★57	0
camelV0	26	0	★30	0	28	0
camelV1	28	0	★38	0	34	0
ivyV0	★26	0	21	0	24	0
jeditV0	41	0	★56	0	52	0
jeditV1	26	0	★33	0	32	0
jeditV2	5	0	★6	0	5	0
log4jV0	★98	0	96	0	95	0
luceneV0	64	0	★65	0	63	0
poiV0	★68	0	65	0	67	0
poiV1	77	0	72	0	★79	0
synapseV0	★64	0	51	0	61	0
velocityV0	34	0	40	0	★41	0
xercesV0	19	0	17	0	★28	0
xercesV1	★88	0	73	0	79	0

Figure 17: precision values comparisons,old runs followed by new runs. In new experiments, 1st, 2nd release are used for training, and test on 3rd release.

Datasets	Tuned_Where	Naive_Where	Tuned_CART	Naive_CART	Tuned_RanFst	Naive_RanFst
antV0	50/ 76.50	1.42	80/ 4.82	0.08	60/ 7.44	0.16
antV1	90/ 235.43	2.31	70/ 5.96	0.08	50/ 8.20	0.20
antV2	70/ 435.70	6.31	70/ 8.71	0.16	60/ 12.81	0.41
camelV0	70/ 528.58	7.89	80/ 13.12	0.18	70/ 17.46	0.30
camelV1	50/ 1498.80	23.87	70/ 18.36	0.24	50/ 22.81	0.71
ivy	50/ 54.83	1.00	50/ 3.04	0.06	50/ 6.40	0.17
jeditV0	50/ 283.42	5.15	80/ 7.58	0.08	50/ 9.03	0.29
jeditV1	50/ 313.84	6.38	70/ 7.04	0.10	50/ 10.25	0.30
jeditV2	50/ 357.03	6.76	70/ 7.92	0.11	50/ 10.30	0.33
log4j	50/ 74.28	1.30	50/ 2.10	0.05	50/ 5.47	0.16
lucene	50/ 162.91	2.78	50/ 3.85	0.08	60/ 10.01	0.24
poiV0	50/ 208.96	3.92	50/ 4.97	0.10	50/ 8.96	0.28
poiV1	70/ 526.41	7.00	60/ 7.18	0.10	60/ 12.95	0.27
synapse	70/ 157.51	1.84	60/ 3.77	0.05	70/ 9.24	0.16
velocity	50/ 154.39	2.58	60/ 4.05	0.05	50/ 6.89	0.19
xercesV0	50/ 103.19	1.99	50/ 5.13	0.08	50/ 9.16	0.21
xercesV1	50/ 686.59	12.61	100/ 13.93	0.15	60/ 14.69	0.38

Figure 14: Time (in seconds) spent on the objective of pd with B stop: tune once+ test once

Datasets	Tuned_Where	Naive_Where	Tuned_CART	Naive_CART	Tuned_RanFst	Naive_RanFst
antV0	50/ 77.16	1.30	60/ 3.47	0.07	50/ 6.54	0.14
antV1	50/ 135.16	2.33	60/ 4.84	0.08	50/ 8.61	0.19
antV2	50/ 368.50	6.23	60/ 6.67	0.14	50/ 11.75	0.40
camelV0	50/ 479.03	8.25	60/ 11.73	0.18	60/ 15.69	0.31
camelV1	60/ 1551.51	26.87	60/ 18.47	0.28	80/ 37.96	0.73
ivy	80/ 83.80	1.01	50/ 3.07	0.07	60/ 7.63	0.17
jeditV0	60/ 349.99	5.95	70/ 8.33	0.11	50/ 13.40	0.30
jeditV1	50/ 442.35	7.48	70/ 8.63	0.11	70/ 15.52	0.37
jeditV2	50/ 360.04	7.90	60/ 8.63	0.12	60/ 14.33	0.38
log4j	60/ 97.23	1.49	70/ 3.31	0.06	80/ 8.79	0.17
lucene	70/ 206.41	2.89	80/ 6.90	0.08	60/ 10.30	0.25
poiV0	70/ 338.44	4.28	60/ 6.47	0.13	60/ 12.45	0.29
poiV1	60/ 468.73	7.95	50/ 7.24	0.14	70/ 17.66	0.30
synapse	50/ 111.30	1.87	60/ 3.85	0.05	60/ 7.82	0.15
velocity	60/ 192.83	2.66	90/ 6.11	0.06	80/ 12.99	0.22
xercesV0	70/ 139.33	2.23	60/ 6.47	0.11	50/ 8.89	0.25
xercesV1	50/ 759.82	13.11	50/ 8.62	0.14	60/ 16.90	0.37

Figure 15: Time (in seconds) spent on the objective of prec with B stop: tune once+ test once

Datasets	Tuned_Where	Naive_Where	Tuned_CART	Naive_CART	Tuned_RanFst	Naive_RanFst
antV0	70/ 95.34	1.18	50/ 3.09	0.06	60/ 6.89	0.13
antV1	50/ 130.69	2.24	50/ 4.47	0.08	60/ 9.03	0.19
antV2	50/ 296.68	6.66	60/ 7.29	0.15	60/ 12.74	0.31
camelV0	60/ 496.70	8.01	60/ 10.43	0.17	70/ 17.32	0.30
camelV1	70/ 1732.99	23.35	60/ 15.93	0.22	60/ 24.71	0.71
ivy	70/ 58.81	1.03	70/ 4.07	0.06	60/ 7.37	0.16
jeditV0	90/ 412.23	5.17	70/ 6.59	0.08	60/ 10.56	0.28
jeditV1	70/ 425.18	6.61	60/ 6.15	0.10	60/ 13.00	0.31
jeditV2	60/ 398.86	6.97	50/ 5.90	0.12	60/ 13.63	0.35
log4j	50/ 68.17	1.37	60/ 2.46	0.04	70/ 6.37	0.15
lucene	50/ 147.78	2.74	50/ 3.79	0.07	60/ 7.98	0.24
poiV0	60/ 251.48	3.90	60/ 5.65	0.09	70/ 11.25	0.30
poiV1	60/ 404.96	7.22	70/ 8.10	0.11	60/ 12.74	0.28
synapse	50/ 95.37	1.80	60/ 3.69	0.06	60/ 7.65	0.14
velocity	60/ 176.80	3.40	50/ 4.14	0.07	50/ 7.36	0.23
xercesV0	60/ 177.48	2.34	50/ 6.53	0.09	50/ 14.91	0.33
xercesV1	50/ 762.28	13.33	70/ 12.84	0.15	60/ 16.63	0.40

Figure 16: Time (in seconds) spent on the objective of F with B stop: tune once+ test once

Data set	WHERE		CART		Random Forests	
	Naive	Tuned	Naive	Tuned	Naive	Tuned
antV0	★39	25	32	31	★39	12
antV1	11	6	40	★45	39	44
antV2	0	★87	44	1	50	51
camelV0	0	28	9	28	★34	30
camelV1	★34	★34	31	32	33	31
ivyV0	27	34	30	★38	35	33
jeditV0	50	56	56	54	★61	59
jeditV1	37	33	36	★49	45	47
jeditV2	4	8	5	★11	9	9
log4jV0	★62	56	47	59	53	43
luceneV0	70	★75	56	★75	73	★75
poiV0	★78	64	74	68	73	69
poiV1	5	★78	21	★78	76	★78
synapseV0	0	2	40	★56	51	55
velocityV0	★51	★51	49	★51	★51	★51
xercesV0	22	20	21	★27	23	20
xercesV1	25	39	18	53	68	★73
antV0	0	0	22	0	★30	0
antV1	★85	0	37	0	39	0
antV2	54	0	45	0	★57	0
camelV0	36	0	39	0	★40	0
camelV1	34	0	39	0	★42	0
ivyV0	★39	0	28	0	36	0
jeditV0	54	0	56	0	★63	0
jeditV1	39	0	44	0	★46	0
jeditV2	9	0	★11	0	8	0
log4jV0	44	0	54	0	★61	0
luceneV0	★73	0	65	0	70	0
poiV0	★50	0	32	0	45	0
poiV1	74	0	68	0	★78	0
synapseV0	★60	0	44	0	52	0
velocityV0	50	0	53	0	★56	0
xercesV0	25	0	19	0	★35	0
xercesV1	21	0	35	0	★42	0

Figure 18: f values comparisons, old runs followed by new runs. In new experiments, 1st, 2nd release are used for training, and test on 3rd release.