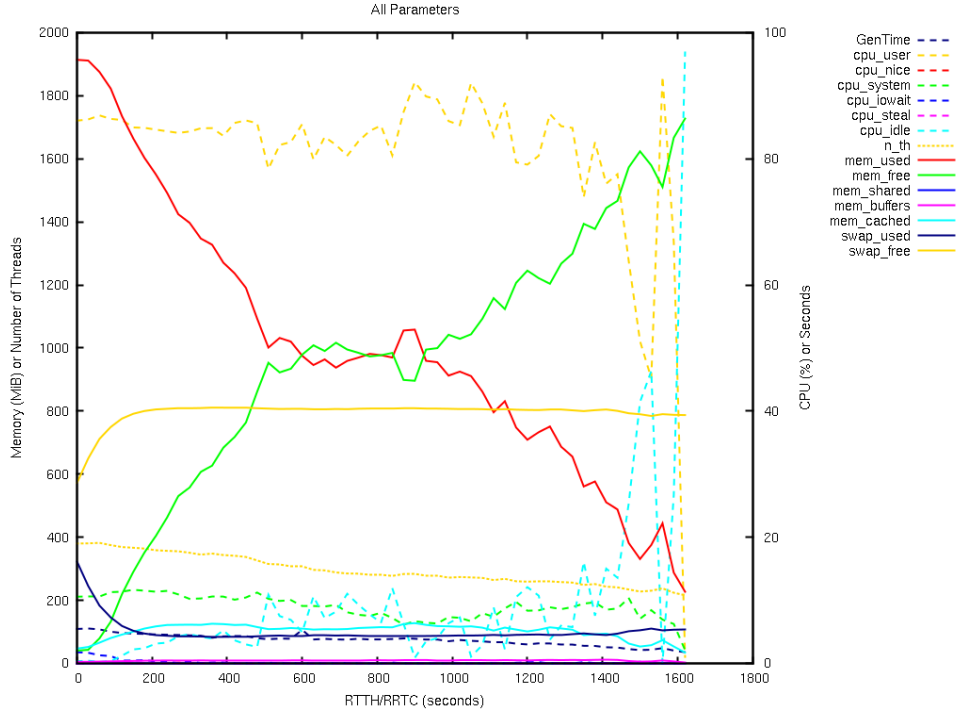# Training Results

Machine Learning Framework

### December 11, 2014

1 System Parameters

The trend of System Parameters as captured by the data collection is here reported:

2 Parameters Selected by Lasso

This Section shows the number of parameters selected by Lasso regularization. Additionally, tables present the specific weights assigned by the regularization process.

λ	# Parameters
$\lambda = 10^0$	23
$\lambda = 10^1$	25
$\lambda = 10^2$	25
$\lambda = 10^3$	25
$\lambda = 10^4$	23
$\lambda = 10^5$	19
$\lambda = 10^6$	15
$\lambda = 10^7$	10
$\lambda = 10^8$	9
$\lambda = 10^9$	6

Number of Parameters Selected by Lasso
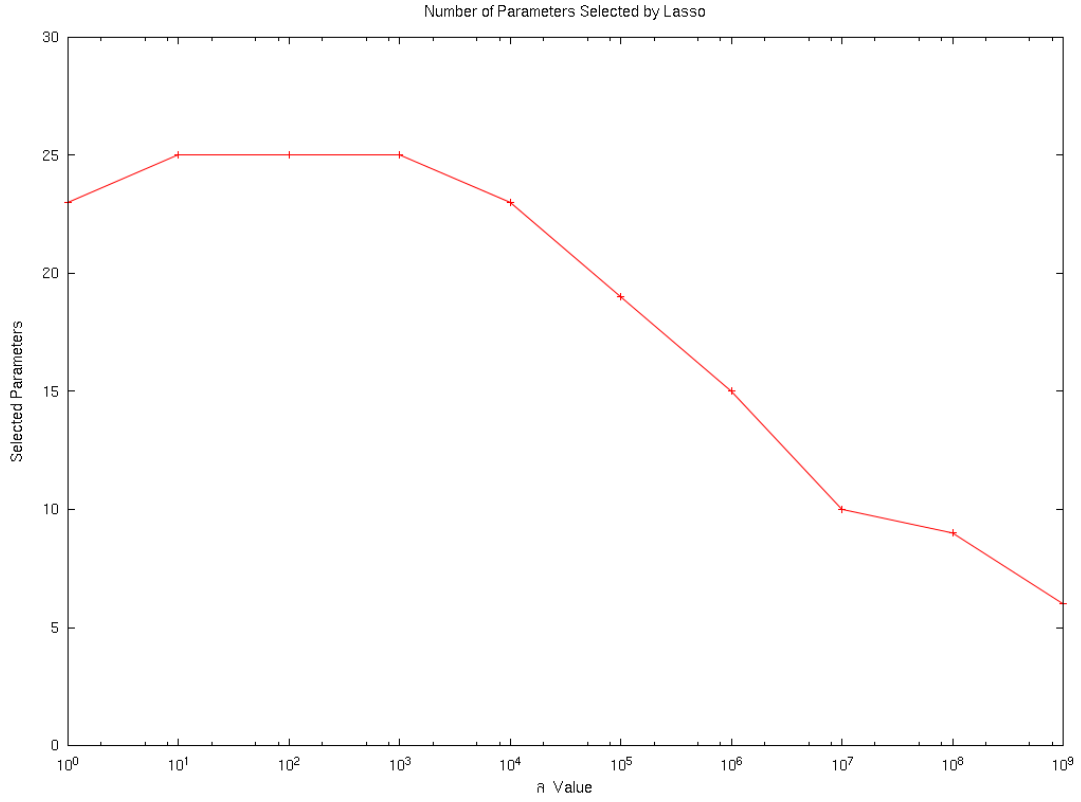
$\lambda = 10^0$

parameter	weight
gen_time	13.077194154074091
n_th_slope	-21.994923984905213
mem_used_slope	-0.003896835087338
mem_buffers_slope	0.000543426630404
mem_cached_slope	-0.000338773086734
swap_used_slope	0.004349199484742
cpu_user_slope	-2.099406808036674
cpu_nice_slope	-10.311089599056476
cpu_system_slope	-9.788773204531411
cpu_iowait_slope	-0.532448955002911
cpu_idle_slope	-6.236734590803698
n_th	-0.501492611404146
mem_used	-6.555948852670959
mem_free	-6.555505408245204
mem_buffers	0.017451192500739
mem_cached	0.000155660466877
swap_used	14.316941166178584
swap_free	14.317529753437986
cpu_user	-6.679837510036281
cpu_nice	-17.588566832583911
cpu_system	-19.703933585255257
cpu_iowait	-5.067314121725085
cpu_idle	-10.918469502053824

$\lambda = 10^1$

2

parameter	weight
gen_time	13.077299733097787
n_th_slope	-21.994464210208058
mem_used_slope	-0.001262071788597
mem_free_slope	0.002634785022062
mem_buffers_slope	0.000543190664417
mem_cached_slope	-0.000338786322463
swap_used_slope	0.000172306165526
swap_free_slope	-0.004177156199197
cpu_user_slope	-2.090384914346590
cpu_nice_slope	-10.283203927330522
cpu_system_slope	-9.779326625864778
cpu_iowait_slope	-0.523143613616289
cpu_idle_slope	-6.227343196373610
n_th	-0.501484792276238
mem_used	-6.550775652170592
mem_free	-6.550332209243444
mem_buffers	0.017451212071472
mem_cached	0.000155663716114
swap_used	14.305643626644139
swap_free	14.306232214566087
cpu_user	-6.673822495000163
cpu_nice	-17.579743525510295
cpu_system	-19.697632347587991
cpu_iowait	-5.061103368033654
cpu_idle	-10.912385600887040

$\lambda = 10^2$

parameter	weight
gen_time	13.077647575497418
n_th_slope	-21.989767607038992
mem_used_slope	-0.003127735233051
mem_free_slope	0.000769168315263
mem_buffers_slope	0.000540216847481
mem_cached_slope	-0.000339178361622
swap_used_slope	0.002697213704695
swap_free_slope	-0.001654902160090
cpu_user_slope	-1.999827267550300
cpu_nice_slope	-10.001688757753293
cpu_system_slope	-9.684398858716124
cpu_iowait_slope	-0.429878349473089
cpu_idle_slope	-6.133480266090081
n_th	-0.501419792494515
mem_used	-6.516269502494544
mem_free	-6.515826073450907
mem_buffers	0.017451357336408
mem_cached	0.000155655916741
swap_used	14.230284863462925
swap_free	14.230873460041311
cpu_user	-6.613553548792910
cpu_nice	-17.491103368696550
cpu_system	-19.634348852268239
cpu_iowait	-4.998979814680190
cpu_idle	-10.851623032468607

$\lambda = 10^3$

parameter	weight
gen_time	13.078130667130546
n_th_slope	-21.941846472528269
mem_used_slope	-0.001699128627569
mem_free_slope	0.002197964124797
mem_buffers_slope	0.000507808015985
mem_cached_slope	-0.000341021552827
swap_used_slope	0.002152574147547
swap_free_slope	-0.002222667257111
cpu_user_slope	-1.429927908435130
cpu_nice_slope	-7.521376586419597
cpu_system_slope	-9.083363388194636
cpu_iowait_slope	0.134828907086072
cpu_idle_slope	-5.530287534552418
n_th	-0.500826971616277
mem_used	-6.199529744012731
mem_free	-6.199086456452998
mem_buffers	0.017452565820228
mem_cached	0.000155536885368
swap_used	13.538551822025868
swap_free	13.539140518852150
cpu_user	-6.079485235395718
cpu_nice	-16.672167311324742
cpu_system	-19.074168661019648
cpu_iowait	-4.455735549707058
cpu_idle	-10.312826902433176

$\lambda = 10^4$

parameter	weight
gen_time	13.065353807364369
n_th_slope	-21.554675255144513
mem_used_slope	-0.001950954840840
mem_free_slope	0.001949490325287
mem_buffers_slope	0.000363006802239
mem_cached_slope	-0.000315992173129
swap_used_slope	0.002282161682561
swap_free_slope	-0.002281667236188
cpu_user_slope	0.170050622879160
cpu_system_slope	-7.373286122507649
cpu_iowait_slope	1.654744717958588
cpu_idle_slope	-3.624519599055949
n_th	-0.494680338496946
mem_used	-2.890818589676904
mem_free	-2.890376681946937
mem_buffers	0.017472264993055
mem_cached	0.000155277199890
swap_used	6.312735569854843
swap_free	6.313326175935761
cpu_user	-1.707701235871729
cpu_nice	-9.990286591090657
cpu_system	-14.506186196696561
cpu_idle	-5.889840408117788

$\lambda = 10^5$

parameter	weight
gen_time	12.158966976232044
n_th_slope	-17.881760262476458
mem_used_slope	-0.001892254462263
mem_free_slope	0.001892240162070
mem_buffers_slope	-0.000777072735794
swap_used_slope	0.002038847258650
swap_free_slope	-0.002038827084764
cpu_user_slope	0.175921351273413
cpu_system_slope	-6.587517412375046
cpu_iowait_slope	0.248981913495916
cpu_idle_slope	-1.515546837047780
n_th	-0.515061387238653
mem_free	0.000433365170104
mem_buffers	0.017568894311829
mem_cached	0.000173039641432
swap_free	0.000531146846394
cpu_user	-1.457666268551556
cpu_system	-14.225848591709662
cpu_idle	-5.468666896699308

$\lambda = 10^6$

parameter	weight
gen_time	5.077852743752413
mem_used_slope	-0.001380882419474
mem_free_slope	0.001380881926018
mem_buffers_slope	-0.007452124226490
mem_cached_slope	0.000738381178203
swap_used_slope	0.002258570580260
swap_free_slope	-0.002258570521081
cpu_user_slope	0.826855000810136
n_th	-0.549976752640693
mem_used	-0.000229228648763
mem_free	0.000149026113255
mem_buffers	0.019981134520878
swap_free	0.000543944628576
cpu_user	3.415102631589505
cpu_system	-8.485828735516449

$\lambda = 10^7$

parameter	weight
mem_used_slope	-0.001621974134002
mem_free_slope	0.001621974147971
mem_cached_slope	0.003378420894753
swap_used_slope	-0.001828570434510
swap_free_slope	0.001828570439847
n_th	-0.465866615479610
mem_free	0.000298294604443
mem_buffers	0.021243063472964
swap_free	0.000134757133453
cpu_user	1.471454160399700

$\lambda = 10^8$

parameter	weight
mem_used_slope	-0.001792106019112
mem_free_slope	0.001792106019412
mem_cached_slope	0.004559306991364
swap_used_slope	-0.001323640582246
swap_free_slope	0.001323640582205
mem_used	-0.000302991753253
mem_buffers	0.022884556617929
swap_used	0.000421011369554
swap_free	0.000768708421149

$\lambda = 10^9$

parameter	weight
mem_used_slope	-0.001349712778213
mem_free_slope	0.001349712778211
swap_used_slope	-0.000602043031343
swap_free_slope	0.000602043031343
mem_free	0.000334016698914
mem_buffers	0.025507773187865

3 Maximum Absolute Prediction Error

Represents the highest error (worst case) encountered during the prediction. It is expressed in seconds.

Algorithm	Error (seconds)
Linear Regression	1029.293
M5P	1038.523
REP Tree	1013.951
SVM	5561.302
SVM2	5559.618
Lasso ($\lambda = 1$)	1179.36458808
Lasso ($\lambda = 10$)	1179.36178165
Lasso ($\lambda = 10^2$)	1179.33647442
Lasso ($\lambda = 10^3$)	1179.08645066
Lasso ($\lambda = 10^4$)	1176.81478187
Lasso ($\lambda = 10^5$)	1171.68082017
Lasso ($\lambda = 10^6$)	1200.60415063
Lasso ($\lambda = 10^7$)	2847.75970936
Lasso ($\lambda = 10^8$)	3113.03534659
Lasso ($\lambda = 10^9$)	2516.47123106

4 Relative Absolute Prediction Error

The relative absolute E_i error is relative to a simple predictor, which is just the average of the actual values. In this case the error is just the total absolute error instead of the total squared error. Thus, the relative absolute error takes the total absolute error and normalizes it by dividing by the total absolute error of the simple predictor.

$$E_i = \frac{\sum_{j=1}^{n} |P_j - T_j|}{\sum_{j=1}^{n} |T_j - \bar{T}|} \tag{1}$$

6

Where P_j is the value predicted for sample j (out of n samples), T_j is the target (i.e., ground truth) value for sample j, and:

$$\bar{T} = \frac{1}{n}\sum_{j=1}^{n} T_j \tag{2}$$

Algorithm	Error (percentage)
Linear Regression	57.701828019%
M5P	34.8564830571%
REP Tree	31.4650074336%
SVM	55.6017117367%
SVM2	55.6012588453%
Lasso ($\lambda = 1$)	165.580653825%
Lasso ($\lambda = 10$)	165.580610152%
Lasso ($\lambda = 10^2$)	165.580234508%
Lasso ($\lambda = 10^3$)	165.576997857%
Lasso ($\lambda = 10^4$)	165.550488453%
Lasso ($\lambda = 10^5$)	165.424465904%
Lasso ($\lambda = 10^6$)	165.108748114%
Lasso ($\lambda = 10^7$)	163.062523904%
Lasso ($\lambda = 10^8$)	163.140486354%
Lasso ($\lambda = 10^9$)	160.381298372%

5   Mean Absolute Error

The Mean Absolute Error is the average of the differences between predicted and real remaining time to failure. Specifically, it is calculated as:

$$\frac{1}{n}\sum_{j=1}^{n} |P_j - T_j| \tag{3}$$

In the following table values are given in seconds

Algorithm	Error (seconds)
Linear Regression	141.336454877
M5P	85.3784345123
REP Tree	77.0712602358
SVM	136.192371919
SVM2	136.191262594
Lasso ($\lambda = 1$)	405.221849085
Lasso ($\lambda = 10$)	405.221742206
Lasso ($\lambda = 10^2$)	405.220822901
Lasso ($\lambda = 10^3$)	405.212901918
Lasso ($\lambda = 10^4$)	405.148026044
Lasso ($\lambda = 10^5$)	404.839613863
Lasso ($\lambda = 10^6$)	404.066964742
Lasso ($\lambda = 10^7$)	399.059285772
Lasso ($\lambda = 10^8$)	399.25008154
Lasso ($\lambda = 10^9$)	392.497582199

6   Soft-Mean Absolute Error

The Soft-Mean Absolute Error is calculated as the Mean Absolute Error except that when the value $|P_j - T_j|$ is below a given threshold it is assumed to be equal to 0.

6.1 Tolerance Threshold: 10%

In this case, if $|P_j - T_j| < 0.1T_j$ then $|P_j - T_j|$ is assumed to be equal 0. In the following table, values are given in seconds.

Algorithm	Error (seconds)
Linear Regression	137.600769453
M5P	79.1823053591
REP Tree	69.8326948553
SVM	132.668273526
SVM2	132.675671061
Lasso ($\lambda = 1$)	405.187785166
Lasso ($\lambda = 10$)	405.187682863
Lasso ($\lambda = 10^2$)	405.186795133
Lasso ($\lambda = 10^3$)	405.178674519
Lasso ($\lambda = 10^4$)	405.124039629
Lasso ($\lambda = 10^5$)	404.823060394
Lasso ($\lambda = 10^6$)	404.041058968
Lasso ($\lambda = 10^7$)	399.023920789
Lasso ($\lambda = 10^8$)	399.240325175
Lasso ($\lambda = 10^9$)	392.469221072

6.2 Tolerance Threshold: 5 minutes (300 seconds)

In this case, if $|P_j - T_j| < 300$ then $|P_j - T_j|$ is assumed to be equal 0. In the following table, values are given in seconds.

Algorithm	Error (seconds)
Linear Regression	51.7908401929
M5P	20.2185858521
REP Tree	14.9696612004
SVM	53.2885028939
SVM2	53.2892733119
Lasso ($\lambda = 1$)	345.477898925
Lasso ($\lambda = 10$)	345.477628507
Lasso ($\lambda = 10^2$)	345.475176064
Lasso ($\lambda = 10^3$)	345.612883
Lasso ($\lambda = 10^4$)	345.595517473
Lasso ($\lambda = 10^5$)	345.335377337
Lasso ($\lambda = 10^6$)	345.475925164
Lasso ($\lambda = 10^7$)	346.709235566
Lasso ($\lambda = 10^8$)	355.474649178
Lasso ($\lambda = 10^9$)	335.287866607

7 Training Time for ML models with WEKA

The Training Time represents the time taken to instantiate a prediction model from the input dataset. It is expressed in seconds.

Algorithm	Training Time (seconds)
linear	0.3
m5p	3.1
REPTree	0.56
svm2	417.41
svm	391.69

8   Validation Time of ML models with WEKA

The Validation Time represents the time needed to validate the accuracy of a model. It is expressed in seconds.

Algorithm	Training Time (seconds)
linear	0.42
m5p	0.36
REPTree	0.55
svm2	0.38
svm	0.39

9   Fitted Models

In this Section we report fitted models for the various machine learning algorithm used by the Framework.

Note that the set of parameters used for training is different: Lasso uses only a subset of parameters, namely the ones associated with non-zero weights in the β vector. On the other hand, the other algorithms rely on the whole set of input parameters.

lasso-lambda-10

lasso-lambda-100

lasso-lambda-1000

lasso-lambda-10000

11

lasso-lambda-100000

lasso-lambda-1000000

12

lasso-lambda-10000000

lasso-lambda-100000000

13

m5p/m5p

Predicted RTTH (seconds)

RTTH (seconds)

reptree/reptree

Predicted RTTH (seconds)

RTTH (seconds)

15

svm/svm

svm2/svm2

neural/neural

Predicted RTTH (seconds) vs RTTH (seconds)