Node Awareness in Elastic Net

L3P3 Center for Open Middleware

Thursday, October 02, 2014

This document summarizes the work done in order to expand the elastic net regularized predictor to include node awareness. This first iteration separates the events data by the node they happened in and then tries to establish elastic net models for each event in each node. The whole process of model creation is described in the file “nen-analytics.R”, found in the private L3P3 Github repository. The models are exported in a variable called “node_separated_model_list”, which is a list of all the nodes in the system, containing each one a list of models for the events that happened on the system.

```r
load('data.R')
head(summary(node_separated_model_list))
```

```
##                         Length Class    Mode  
## aswpeuib01.sceu.corp "24"   "-none-" "list"
## aswpeuib02.sceu.corp "28"   "-none-" "list"
## aswpeuib03.sceu.corp "24"   "-none-" "list"
## aswpeuib04.sceu.corp "26"   "-none-" "list"
## aswpeuid01.sceu.corp "28"   "-none-" "list"
## aswpeuid02.sceu.corp "31"   "-none-" "list"
```

Our objective is now to study the performance of this models. To do so, we’re going to create a matrix where the rows are the system nodes and the columns are the events that happened on the system. The coefficients of the matrix will be the f-score of each model in each node.

```r
events<-extract_unique_events(node_separated_model_list)
model_matrix<-matrix(nrow=length(node_separated_model_list),ncol=length(events))
colnames(model_matrix)<-events
rownames(model_matrix)<-names(node_separated_model_list)
for (i in 1:nrow(model_matrix)){
  for (j in 1:ncol(model_matrix)){
    model_index<-which(names(node_separated_model_list[[i]])==colnames(model_matrix)[j])
    if (length(model_index)==0){ #The event doesn't happen in that node
      model_matrix[i,j]<-NaN
      }
    else{
      model_matrix[i,j]<-node_separated_model_list[[i]][[model_index]]$fscore
    }
  }
}
model_matrix[1:5,1:5]
```

```
##                      NO.EVENTS 66305  66306 68912  68918
## aswpeuib01.sceu.corp    0.4444     1    NaN   NaN    NaN
## aswpeuib02.sceu.corp    0.2000     1 1.0000     1 1.0000
## aswpeuib03.sceu.corp    0.5455   NaN 0.6667     1 0.6667
## aswpeuib04.sceu.corp    0.4000     1    NaN     1    NaN
## aswpeuid01.sceu.corp    0.6667     1    NaN   NaN    NaN
```

Let’s now extract some insightful analytics.

- First, let’s see how many models were correctly created out of the possible amount for each node and the average fscore for each node’s models:

```
##                             Events Models Percentage Average fscore
## aswpeuib01.sceu.corp            31      9     29.03%         0.8526
## aswpeuib02.sceu.corp            34     14     41.18%         0.8662
## aswpeuib03.sceu.corp            28     11     39.29%         0.8181
## aswpeuib04.sceu.corp            32      6     18.75%         0.8134
## aswpeuid01.sceu.corp            32      5     15.62%         0.9253
## aswpeuid02.sceu.corp            34      7     20.59%         0.8770
## aswpeuid03.sceu.corp            38      5     13.16%         0.8089
## aswpeuid04.sceu.corp            38      6     15.79%         0.7715
## aswpeuin01.sceu.corp            36     13     36.11%         0.8942
## aswpeuin02.sceu.corp            34      9     26.47%         0.8190
## aswpeuin03.sceu.corp            37      8     21.62%         0.8150
## aswpeuin04.sceu.corp            35     14        40%         0.8366
## aswpeuin05.sceu.corp            33     12     36.36%         0.8745
## aswpeuin06.sceu.corp            31     12     38.71%         0.8879
## aswpeuin07.sceu.corp            27     11     40.74%         0.8306
## aswpeuin08.sceu.corp            31     12     38.71%         0.8674
## aswpeuin09.sceu.corp            28     12     42.86%         0.8636
## aswpeuin10.sceu.corp            28      8     28.57%         0.7835
## aswpeuin11.sceu.corp            24      6        25%         0.7906
## aswpeuin12.sceu.corp            28      7        25%         0.7734
## ASWPEUIN13.sceu.corp            27     11     40.74%         0.8323
## ASWPEUIN14.sceu.corp            22     13     59.09%         0.8155
## ASWPEUIN15.sceu.corp            20      5        25%         0.7062
## ASWPEUIN16.sceu.corp            18     10     55.56%         0.8766
## ASWPEUIN17.sceu.corp            21     11     52.38%         0.8501
## ASWPEUIN18.sceu.corp            20     12        60%         0.8202
## ASWPEUIN19.sceu.corp            31     10     32.26%         0.9048
## ASWPEUIN20.sceu.corp            20     10        50%         0.8581
## dnsautpeuid01-adm.sceu.corp      8      1      12.5%         0.6667
## dnsautpeuid02-adm.sceu.corp      6      1     16.67%         0.5000
## dnsrespeuin01-adm.sceu.corp     11      1      9.09%         0.6667
## dnsrespeuin02-adm.sceu.corp     13      1      7.69%         0.8000
## dswpeuin01.sceu.corp            44     18     40.91%         0.8995
## dswpeuin02.sceu.corp            43     12     27.91%         0.8783
## dswpeuin03.sceu.corp            34     10     29.41%         0.7093
## dswpeuin04.sceu.corp            35     21        60%         0.8224
## dswpeuwt01.sceu.corp            45      9        20%         0.9431
## dswpeuwt02.sceu.corp            40     11      27.5%         0.7728
## dswpeuwt03.sceu.corp            39      5     12.82%         0.7631
## dswpeuwt04.sceu.corp            39      5     12.82%         0.6644
## fwpeuib01.sceu.corp              7      3     42.86%         0.8833
## fwpeuib02.sceu.corp              8      5      62.5%         0.8638
## fwpeuid01.sceu.corp              8      1      12.5%         0.7442
## fwpeuid02.sceu.corp             16      1      6.25%         0.6522
## FWPEUIN01                       10      2        20%         0.7000
## FWPEUIN02                       11      1      9.09%         1.0000
## lbpeuib01.sceu.corp             22      5     22.73%         0.7076
## lbpeuib02.sceu.corp             23      3     13.04%         0.8000
## lbpeuid01.sceu.corp             23     10     43.48%         0.6179
## lbpeuid02.sceu.corp             23      8     34.78%         0.5908
## lbpeuin01.sceu.corp            26      10     38.46%          0.8712
## lbpeuin02.sceu.corp            26       7     26.92%          0.8548
```

Even though most of the events are not captured, the ones that are modelled show good performance.

- Now, we'll study the number of nodes an event appears in, the amount of models created for that certain event and the average fscore for that model:

Critical Severity

```
##          Total.Nodes Created.Models Percentage Average.fscore
## 96731154           2              1        50%         0.6667
## 68917             26              0         0%            NaN
## 69481             24              3      12.5%         0.8974
```

Major Severity

```
##            Total.Nodes Created.Models Percentage Average.fscore
## 62324754             6              4     66.67%         0.5876
## 12845059             2              0         0%            NaN
## 2228225             24             12        50%         0.7762
## 69379                7              0         0%            NaN
## 96731153            30              2      6.67%         1.0000
## 69034                1              0         0%            NaN
## 4293918809           3              1     33.33%         1.0000
## 69489               36             19     52.78%         0.9087
## 4293918944           6              0         0%            NaN
```

Minor Severity

```
##          Total.Nodes Created.Models Percentage Average.fscore
## 62324756           6              4     66.67%         0.5661
```

Blank Severity

```
##            Total.Nodes Created.Models Percentage Average.fscore
## 62324755             6              5     83.33%         0.6027
## 62324757             6              6       100%         0.7652
## 62324754             6              4     66.67%         0.5876
## 1116676             38              0         0%            NaN
## 1116675             52              0         0%            NaN
## 2228482             21             10     47.62%         0.8511
## 2228226             24             12        50%         0.8860
## 8523866             22              2      9.09%         0.9000
## 8523849             23              1      4.35%         0.8000
## 2162726             31              6     19.35%         0.9792
## 8523870             22             11        50%         1.0000
## 8523803             22             10     45.45%         1.0000
## 8523778             22              3     13.64%         1.0000
## 2162714             23              1      4.35%         0.7619
## 8523816             20              0         0%            NaN
## 8523812             22              0         0%            NaN
## 8523798             36              0         0%            NaN
## 62324771             6              2     33.33%         0.9000
## 2162711             34             26     76.47%         0.8037
## 851970              15              7     46.67%         0.8336
## 2163911             28             20     71.43%         0.8658
## 2228481             10              4        40%         0.8731
## 69378                8              0         0%            NaN
## 96731145            32              2      6.25%         0.7857
## 67335                6              0         0%            NaN
## 68876                6              0         0%            NaN
## 68864                6              0         0%            NaN
## 39714818             4              0         0%            NaN
## 4293918944           6              0         0%            NaN
## 69382                6              1     16.67%         1.0000
## 66305               52             16     30.77%         0.9688
## 69386                6              1     16.67%         1.0000
## 69128               36              1      2.78%         1.0000
## 67348               14              1      7.14%         0.6667
## 88211715             3              0         0%            NaN
## 88211714             3              1     33.33%         1.0000
## 12845061            12              1      8.33%         0.9474
## 65565               11              1      9.09%         1.0000
## 73137                2              0         0%            NaN
## 68912               47             27     57.45%         0.9333
## 66306               47             15     31.91%         0.8778
## 96731154             2              1        50%         0.6667
## 2162715              4              0         0%            NaN
## 62324759             6              0         0%            NaN
## 62324758             6              0         0%            NaN
## 69547                3              0         0%            NaN
## 12845059             2              0         0%            NaN
## 62324756             6              4     66.67%         0.5661
## 65872                1              0         0%            NaN
## 65564                1              0         0%            NaN
## 8523811              1              0         0%            NaN
## 8523799              1              0         0%            NaN
## 18613386             4              4       100%         0.9819
## 69448               36             31     86.11%         0.8094
## 69450               34             30     88.24%         0.7459
## 69449               24              8     33.33%         0.9722
## 69447               24             11     45.83%         0.9710
## 68918               40             21      52.5%         0.8748
## 69489               36             19     52.78%         0.9087
## 69482               30             16     53.33%         0.9151
## 69490                9              0         0%            NaN
## 69480                9              2     22.22%         1.0000
## 67349               12              1      8.33%         1.0000
## 1116678             28              0         0%            NaN
## 8523864             28              0         0%            NaN
## 66049               28              0         0%            NaN
## 8523827             28              0         0%            NaN
## 66075               28              0         0%            NaN
## 8523824                 1              0          0%          NaN
## 8523823                 1              0          0%          NaN
## 65539                   1              0          0%          NaN
## 2228225                24             12         50%       0.7762
## 39714820                1              0          0%          NaN
## 67329                  14              7         50%       0.7821
## 67330                  14              5     35.71%       0.9170
## 67334                   4              0          0%          NaN
```

Again, it looks like the performance here follows an "all-or-nothing" pattern: there are a lot of nods where patterns where not captured, but when they are the performance is really high.

The total possible models are 1378, whereas the amount of created models is 420, making a percentage of 30.48% modelled events.

Taking into account the performance of the created models, most of them could be incorporated into our predictor proof of concept black box.