Hierarchical Multi-Label Classification Using Local Neural Networks

Cerria, R., Barrosa, R. C., & de Carvalhoa, A. C.
Journal of Computer and Systems Sciences, 2014

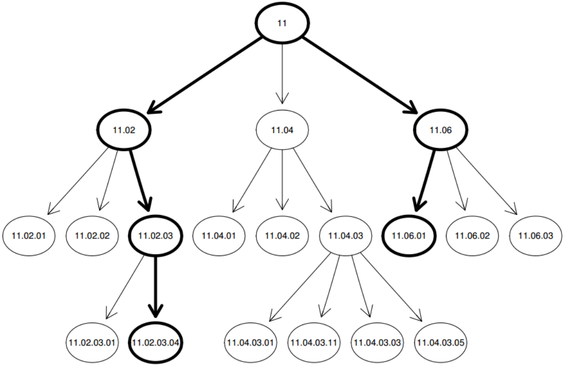
Speaker : Semin Choi

Department of Statistics, Seoul National University, South Korea

May 4, 2017

Hierarchical Multi-Label Classification

- The classes involved in the classification task are hierarchically structured.
- Each example may simultaneously belong to more than one class in each hierarchical level.



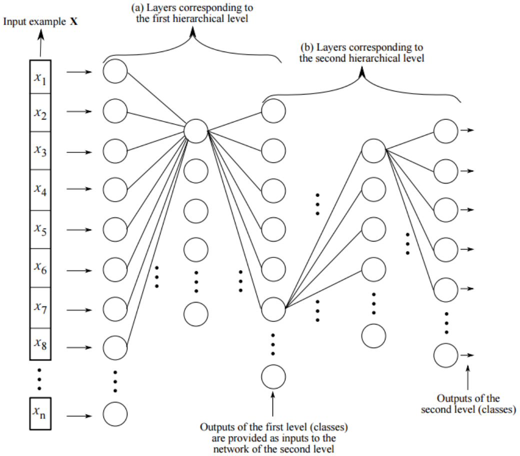
Two main approaches for HMC problems

- Local Methods
  - Local Classifier per Node (LCN)
    - The LCN strategy trains one binary classifier for each class of the hierarchy.
  - Local Classifier per Parent Node (LCPN)
    - The LCPN strategy trains, for each internal class, a multi-class classifier to distinguish between its subclasses.
  - Local Classifier per Level (LCL)
    - The LCL strategy trains one multi-class classifier for each hierarchical level, where each classifier is responsible for the prediction in its associated level.

Two main approaches for HMC problems

- Global Methods
  - The global approach induces only one classifier using all classes of the hierarchy at once.
  - After training the classifier, the classification of a new example occurs in just one step.
  - Since global-based methods induce just one classifier to consider the specificities of the classification problem, they do not use conventional classification algorithms, unless these are adapted to consider the hierarchy of classes.

Hierarchical Multi-Label Classification with Local Multi-Layer Perceptron (HMC-LMLP)



Training Process

- The network architecture has one hidden layer and one output layer(classes of the first level).
- We employed the Back-propagation and Resilient back-propagation algorithms.
- When the training of the MLP network associated with the first hierarchical level is finished, then a second neural network is associated with the next level of the hierarchy.
- The only difference is that the inputs of this network are now the outputs provided by the previous trained MLP.
- When training a MLP network for a specific hierarchical level, the MLP networks associated with the previous levels are not re-trained.

Final Predictions

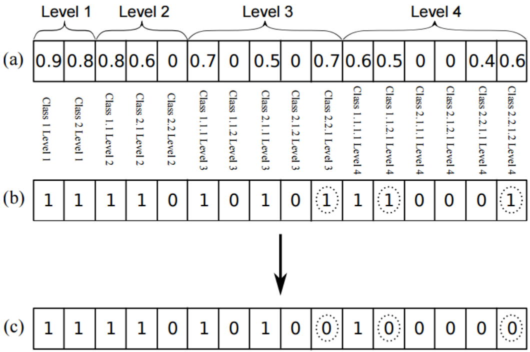


Figure 3: Example of the class-predicted vector provided by HMC-LMLP. (a) outputs of the neurons; (b) prediction after applying a threshold value of 0.5; (c) final classification after correcting inconsistencies.

Evaluation Measures

- We used precision-recall curves as the evaluation measure for the methods.
- We used two PR curve variations to compare the investigated methods :
  - The area under the average PR-curve $(AU(\overline{PRC}))$.
  - The weighted average of the areas under the individual (per class) PR curves $(\overline{AUPRC_w})$.

Experimental Analysis

Table 4: Comparisons between HMC-LMLP and Clus-variations.

Dataset	Bp	Rprop	Clus-HMC	Clus-HSC	Clus-SC
		$AU(\overline{PRC})$ values			
Cellcycle	0.185 ± 0.001	0.183 ± 0.001	0.172	0.111	0.106
Church	0.164 ± 0.001	0.163 ± 0.001	0.170	0.131	0.128
Derisi	0.170 ± 0.001	0.169 ± 0.001	0.175	0.094	0.089
Eisen	0.208 ± 0.001	0.206 ± 0.002	0.204	0.127	0.132
Gasch1	0.196 ± 0.001	0.199 ± 0.001	0.205	0.106	0.104
Gasch2	0.184 ± 0.001	0.183 ± 0.003	0.195	0.121	0.119
Pheno	0.159 ± 0.001	0.155 ± 0.002	0.160	0.152	0.149
Spo	0.172 ± 0.001	0.168 ± 0.001	0.186	0.103	0.098
Expr	0.196 ± 0.003	0.200 ± 0.002	0.210	0.127	0.123
Seq	0.195 ± 0.003	0.203 ± 0.002	0.211	0.091	0.095
Hom	0.195 ± 0.002	0.182 ± 0.004	0.254	0.155	0.153
Struc	0.154 ± 0.0008	0.152 ± 0.0009	0.181	0.118	0.114
Average	0.181	0.180	0.194	0.120	0.118
		$\overline{AUPRC_w}$ values			
Cellcycle	0.145 ± 0.001	0.144 ± 0.002	0.142	0.146	0.146
Church	0.118 ± 0.0007	0.118 ± 0.002	0.129	0.127	0.128
Derisi	0.127 ± 0.0009	0.127 ± 0.001	0.137	0.125	0.122
Eisen	0.163 ± 0.001	0.163 ± 0.003	0.183	0.169	0.173
Gasch1	0.157 ± 0.002	0.158 ± 0.002	0.176	0.154	0.153
Gasch2	0.142 ± 0.001	0.144 ± 0.006	0.156	0.148	0.147
Pheno	0.114 ± 0.0009	0.111 ± 0.002	0.124	0.125	0.127
Spo	0.129 ± 0.001	0.125 ± 0.002	0.153	0.139	0.139
Expr	0.167 ± 0.003	0.165 ± 0.002	0.179	0.167	0.167
Seq	0.166 ± 0.002	0.168 ± 0.002	0.183	0.151	0.154
Hom	0.159 ± 0.002	0.146 ± 0.006	0.240	0.205	0.205
Struc	0.112 ± 0.0008	0.108 ± 0.001	0.161	0.152	0.152
Average	0.142	0.140	0.164	0.151	0.151