

ニューロデータマイニングによるゲノムデータベースからの知識発見

●林 陽一

明治大学理工学部

背景と目的

背景としてはヒトゲノムデータベースから何らかの意味のあるルールを明示的に抽出するには、いったいどのような方法が現実的であろうか。というのが出発点である。

検討結果

今現在、プログラム化可能なニューラルネットワークからのルール抽出(rule extraction)アルゴリズムを用いて実際の肝胆道系疾患のデータベース(536人分)からルールの抽出が実際に現実的な時間で行えるかを、Sun Ultra 80(4cpu)1GB memory(速度約10Gflops)で行ってみた。

対象にしたのは、以下の6つのアルゴリズムであり、上記データベースでのテスト(未学習)データを対象に時間と精度を比較する。

FERNN	時間	4:01(分)	精度	82.3%
NeuroLinear	時間	45:34(分)	精度	90.2%
Mlp21n	時間	5:52(分)	精度	68.2%
Abe's Method	時間	0.08(秒)	精度	77.9%
BIO-RE	時間	123.54(秒)	精度	70.78%
Partial-RE	時間	110.20(秒)	精度	76.68%

上記を見るとAbe's Methodが群を抜いて高速なことがわかるが、この方法は、いわゆるrule extractionを行っているとは言い難く、ルールは得られるものの、前件部がもしn次元直方体に入っているなら等の直感的

でないものになってしまう。上記の中ではPartial-RE、BIO-REも良さそうに見えるが、アルゴリズムの後処理(カルノー図をルールに変換する)が難しく、結局FERNNが総合的に見て良好である。

また、上記のアルゴリズムのどれが容易に並列化(Parallelization)かという点が重要だが、今年度は研究時間と超並列計算機の手当ができなかったので今後の検討課題としたい。

成果公表リスト

1. Duch, W., Adamczak, R., Grabczewski, K., Zal, G and Hayashi, Y. Fuzzy and Crisp logical rule extraction methods in application to medical data, Fuzzy Systems in Medicine (Szczepaniak, P.s. at al. Eds.), Springer Verlag (Studies in Fuzziness and Sot Computing, Vol.41 pp.593-616 (2000))
2. Mitra, S. and Hayashi, Y., Neuro-Fuzzy Rule Generation: Survey in Soft Computing Framework, IEEE Trans. on Neural Networks, Vol.11, No.3 (2000) pp.748-768
3. Hayashi, Y., Setiono, R. and Yoshida, K., A comparison between two neural network rule extraction techniques for the diagnosis of hepatobiliary disorders, Artificial Intelligence in Medicine Vol.20, No.3, pp.205-216 (2000)
4. Duch, W. and Hayashi, Y., Computational intelligence methods and data understanding, In: Quo Vadis computational Intelligence? New trends and approaches in computational intelligence. (P. Sincak, J. Vascak, Springer Eds.) Studies in Fuzziness and Soft Computing, Vol. 54, pp. 256-270 Springer Verlag (2000)
5. Duch, W., Rafal, A and Hayashi, Y., Eliminators and Classifiers, Proceedings of 7th International Conference on Neural Information Processing (ICONIP' 2000), November 14-18, 2000, Taejon, Korea, pp.1029-1034