# On the Validation of the DIMLP Neural Network

**Guido Bologna**

Swiss Institute of Bioinformatics
Rue Michel Servet 1, CH-1211 Geneva, Switzerland
Guido.Bologna@isb-sib.ch

**Abstract**

Rules are extracted from the DIMLP neural network in polynomial time with respect to the size of the classification problem and the size of the network. With rules is possible to ask how well do inferences made compare with knowledge and heuristics of experts. Although fidelity of generated rules from the training set is 100%, perfect fidelity on new unknown data samples is not guaranteed. In this work we introduce a local dynamic algorithm that makes rules consistent with new unknown cases. The presented method is computationally tractable and produces small changes in a rulebase.

## Introduction

When artificial neural networks are used for automatic classification of data samples, a major drawback is that "knowledge" embedded therein is cryptically coded as a large number of weights and activation values. The lack of neural network validation tools is often one of the reasons limiting their use in practice, especially in the context of critical domain applications. Generating symbolic rules from a neural network allows us to ask how well do inferences made compare with knowledge and heuristics of experts in the field. Therefore, rules embedded within a neural network are important for validation purposes.

The degree of matching between neural network responses and classifications by extracted rules is denoted as *fidelity*. Generally, rule extraction techniques do not generate rules with 100% fidelity on unknown cases (Andrews, Diederich, and Tickle 1995) (Duch, Adamczak, and Grabczewski 2001). However, to validate a neural network in a critical domain application, we advocate the view that fidelity on training cases should be equal to 100%, and as high as possible on unknown cases.

In previous work we presented the *Discretized Interpretable Multi Layer Perceptrons* (DIMLP) (Bologna 2000a), (Bologna 2000b). Symbolic rules are generated from a DIMLP network in polynomial time with respect to the dimensionality of the problem, the number of examples, and the size of the network. Moreover, fidelity of rules is 100% with respect to the data samples used during the training phase. Finally, continuous inputs do not need to be transformed to binary values, as is done in the majority of rule extraction techniques (Duch, Adamczak, and Grabczewski 2001).

In this work, with 17 medical problems of the public domain we show that DIMLP networks are as accurate as standard multi-layer perceptrons (MLP). Moreover, the average predictive accuracy of rules is even higher when rules and network agree. For those unknown cases in which rules and network do not agree we prone the strategy of *dynamic validation*. This is achieved with the use of a local dynamic method that makes rules consistent with new unknown cases.



## A Method for Dynamic Validation

Rules are "If-Then" expressions based on boolean logic. Generally, any data sample is able to activate one or more rules. Fidelity of rules generated from DIMLP networks is 100% for all data samples used during the rule extraction phase. Note that perfect fidelity on unknown cases is not guaranteed, because several regions in the space of attributes may be unexplored.

When a new and unknown case is classified and the class of an activated rule does not correspond to the classification of the network, the current rulebase should be updated in order to make rules consistent with neural network responses. At this point two problems must be addressed. First, if the current rulebase has been validated by an expert of the field of application, a new validation procedure must be carried out with the updated rulebase. Therefore, if too many changes have been made, again, a great effort will be put in the validation procedure. The second problem is that the use of the rule extraction algorithm, though polynomial in its nature, may be time consuming. As a consequence, we propose a local method that produces small changes in a rulebase (thus retaining its main validated components), and that scales polynomially with the number of new created antecedents. Our method is given as:

1. Create a new rule covering the new case with the use of the antecedents of the inconsistent rule.
2. Create new rules covering the old rule without the new rule.
3. Prune rule antecedents (if possible).

4. Prune rules (if possible).

A two-dimensional example is illustrated in figure 1 by an inconsistent rule for a classification problem with two classes. A new unknown case indicated by an arrow is inconsistent with Rule $R_1$ given as: "IF $(x_1 \geq a)$ AND $(x_2 \geq d)$ then BLACK_CIRCLE" (A). In (B) the inconsistent case classified as WHITE_CIRCLE by the neural network and covered by $R_1$ of class BLACK_CIRCLE is used to create a new rule $R_2$ given as "IF $(b \leq x_1 \leq c)$ AND $(e \leq x_2 \leq f)$ then WHITE_CIRCLE". Parts (C), (D), (E), and (F) of figure 1 correspond to the creation of 4 new rules denoted as $R_{11}, R_{12}, R_{13}$, and $R_{14}$. They cover the surface of $R_1$ without $R_2$ (cf. step 2). Note that the new rules are

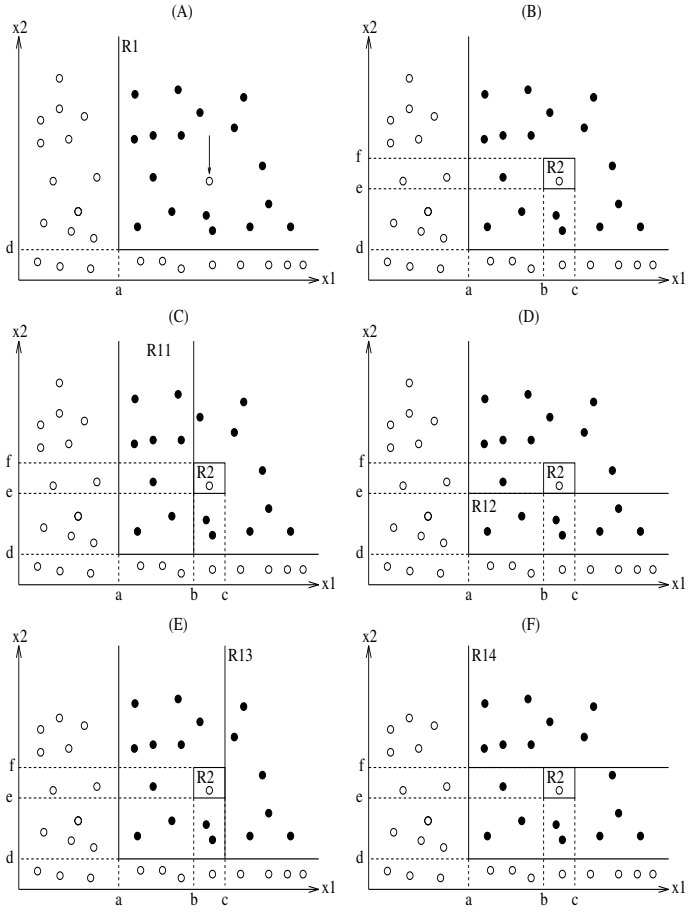


Figure 1: An example of the method used to solve the inconsistency of a data sample.

similar to $R_1$. In fact $R_{11}, R_{12}, R_{13}$, and $R_{14}$ are given as:

- "IF $(a \leq x_1 \leq b)$ AND $(x_2 \geq d)$ then BLACK_CIRCLE";
- "IF $(x_1 \geq a)$ AND $(d \leq x_2 \leq e)$ then BLACK_CIRCLE";
- "IF $(x_1 \geq c)$ AND $(x_2 \geq d)$ then BLACK_CIRCLE";
- "IF $(x_1 \geq a)$ AND $(x_2 \geq f)$ then BLACK_CIRCLE";

An expert who has validated $R_1$ would be able to know whether the new generated rules are relevant or not. Generally, the number of created rules depends on the number of antecedents represented in the inconsistent rule. For instance, here we generated $4 + 1 = 5$ rules. With $n$ antecedents, our algorithm generates $2n + 1$ rules.

## Experiments

Seventeen data sets corresponding to 17 medical diagnosis and medical prognosis problems were selected from the University of California public domain (Blake, and Merz 1998). The purpose of the experiments was to compare our network model to standard multi-layer perceptrons and to demonstrate that the fidelity of extracted rules is reasonably high in order to minimize the number of dynamic validations by experts.

### Methodology

Table 1 gives the characteristics of the data sets, as well as DIMLP and MLP neural architectures. The number of neurons in the hidden layers was based on the heuristic that the number of weights must be less than the number of examples, and three neurons in the second hidden layer being a minimum. Neural networks were trained with default learning parameters (Bologna 2000b).

| DATA SET | CASES | ATTRIBUTES | CLASSES |
|---|---|---|---|
| ARRHYTHMIA | 452 | 279 | 2 |
| BREAST-CANCER | 569 | 30 | 2 |
| BREAST-W | 699 | 9 | 2 |
| DERMATOLOGY | 366 | 34 | 6 |
| ECHOCARDIOGRAM | 131 | 7 | 2 |
| HEART-C | 303 | 13 | 2 |
| HEART-H | 294 | 13 | 2 |
| HEART-STATLOG | 270 | 13 | 2 |
| HEPATITIS | 155 | 19 | 2 |
| HORSE-COLIC | 368 | 22 | 2 |
| HYPOTHYROID | 3772 | 29 | 4 |
| LIVER-DISORDERS | 345 | 6 | 2 |
| LYMPHOGRAPHY | 148 | 15 | 4 |
| PIMA-INDIANS | 768 | 8 | 2 |
| PRIMARY-TUMOUR | 339 | 17 | 22 |
| SICK | 3772 | 29 | 2 |
| THYROID-BENCHMARK | 7200 | 21 | 3 |

Table 1: Data sets and their characteristics.

Our results were based on the average predictive accuracy calculated after 10 repetitions of 10-fold cross-validation. Further, the training phase of MLP and DIMLP networks was stopped according to the minimal error measured on an independent validation set. For each cross-validation trial the proportions of training sets, validation sets, and testing sets were $8/10$, $1/10$ and $1/10$, respectively.

### Results

Average predictive accuracy results are shown in table 2. It turned out that the average predictive accuracy of DIMLPs and MLPs on the 17 classification problems is very close.

From left to right in table 3 are indicated the data set, the fidelity of DIMLP rules on the testing set, the accuracy of DIMLP rules when rules were in agreement with DIMLP responses, and the number of antecedents per rule set. It is worth noting that fidelity was more than 99% for the problems with several thousands of data samples and often more than 90% in the others. When DIMLP rule classifications were equal to DIMLP responses, the average predictive accuracy was significantly higher than the predictive accuracy on all testing cases.

| DATA SET | DIMLP (%) | MLP (%) |
|---|---|---|
| ARRHYTHMIA | 75.7±1.3 | 75.1±0.8 |
| BREAST-CANCER | 96.5±0.5 | 97.3±0.5 |
| BREAST-W | 96.5±0.3 | 96.5±0.4 |
| DERMATOLOGY | 95.7±0.5 | 94.7±1.1 |
| ECHOCARDIOGRAM | 68.5±3.9 | 65.7±2.9 |
| HEART-C | 80.6±1.4 | 80.1±1.4 |
| HEART-H | 79.6±0.9 | 77.8±1.5 |
| HEART-STATLOG | 82.1±2.0 | 82.9±1.6 |
| HEPATITIS | 79.1±2.8 | 79.7±2.0 |
| HORSE-COLIC | 82.1±1.5 | 80.6±1.4 |
| HYPOTHYROID | 99.1±0.1 | 98.2±0.2 |
| LIVER-DISORDERS | 70.1±1.9 | 70.2±1.5 |
| LYMPHOGRAPHY | 80.4±2.0 | 81.6±2.1 |
| PIMA-INDIANS | 75.4±1.1 | 75.5±0.6 |
| PRIMARY-TUMOUR | 40.4±2.3 | 45.0±1.2 |
| SICK | 98.1±0.1 | 97.6±0.2 |
| THYROID-BENCHMARK | 98.4±0.1 | 97.9±0.1 |
| **AVERAGE** | 82.3 | 82.1 |

Table 2: Average predictive accuracy.

| DATA SET | FID. (%) | ACC. (%) | ANT. |
|---|---|---|---|
| ARRHYTHMIA | 88.4±1.7 | 77.3±1.0 | 241.7 |
| BREAST-CANCER | 97.3±0.5 | 97.2±0.4 | 51.1 |
| BREAST-W | 98.5±0.5 | 97.0±0.2 | 27.4 |
| DERMATOLOGY | 95.8±1.1 | 97.4±0.5 | 71.0 |
| ECHOCARDIOGRAM | 96.0±2.0 | 69.4±3.8 | 13.5 |
| HEART-C | 94.1±1.3 | 82.4±1.3 | 74.3 |
| HEART-H | 95.9±1.6 | 80.9±1.0 | 44.5 |
| HEART-STATLOG | 93.9±1.4 | 84.0±1.8 | 69.2 |
| HEPATITIS | 93.9±1.5 | 80.4±2.9 | 18.1 |
| HORSE-COLIC | 96.1±1.3 | 83.4±1.4 | 59.0 |
| HYPOTHYROID | 99.8±0.1 | 99.3±0.1 | 42.4 |
| LIVER-DISORDER | 91.5±1.6 | 71.2±1.6 | 101.7 |
| LYMPHOGRAPHY | 93.3±1.1 | 82.4±2.5 | 41.1 |
| PIMA-INDIANS | 95.8±0.8 | 76.4±1.0 | 132.5 |
| PRIMARY-TUMOUR | 87.8±1.4 | 42.9±2.4 | 256.9 |
| SICK | 99.7±0.1 | 98.3±0.1 | 61.0 |
| THYROID-BENCHMARK | 99.4±0.1 | 98.8±0.1 | 185.4 |
| **AVERAGE** | 95.1 | 83.5 | 87.7 |

Table 3: Average predictive accuracies of rules and average number of antecedents per rule set (see text).

## Discussion and Conclusion

With the help of rules a DIMLP network is able to be validated by experts. Experiments on 17 classification problems showed that DIMLPs are as accurate as standard multi-layer perceptrons, and that the fidelity of generated rules is reasonably high. Moreover, when rules agreed with network responses on testing examples, the predictive accuracy was even better than the predictive accuracy measured on all testing examples. For inconsistent unknown cases we presented a method that generates rules with similar antecedents compared to old inconsistent rules; thus facilitating dynamic validation by experts. In our view of neural network validation, as long as new unknown cases are presented to a network and rules are inconsistent, a rulebase is a dynamic knowledge support that evolves, hopefully with increasing fidelity.

## References


### Book with Multiple Authors

Bologna, G. Symbolic Rule Extraction from the DIMLP Neural Network. 2000a. In: Wermter, S., and Sun, R. eds., *Neural Hybrid Systems*. Springer Verlag: 240–254.

### Journal Article

Andrews, R., Diederich, J., and Tickle, A.B. 1995. Survey and Critique of Techniques for Extracting Rules from Trained Artificial Neural Networks. *Knowl Based Syst* 8(6): 373–389.

Duch, W., Adamczak, R., and Grabczewski, K. 2001. A New Methodology of Extraction, Optimization and Application of Crisp and Fuzzy Logical Rules. *IEEE Transactions on Neural Networks* 12 (2): 277–306.

### Proceedings Paper Published by a Society

Bologna, G. 2000b. Rule Extraction from a Multi-Layer Perceptron with Staircase Activation Functions. In Proceedings of the International Joint Conference on Neural Networks, (3) 419–424.

### University Technical Report

Blake, C.L., and Merz, C.J. 1998. UCI Repository of Machine Learning Databases (http://www.ics.uci.edu/m̃learn/MLRepository.html). Department of Information and Computer Science. University of California.