Feasibility of Classification-Based Algorithms for the AGATA Pulseshape Analysis

M. Schlarb, R. Krücken, and R. Gernhäuser

The AGATA detector array will consist of up to 192 36-fold segmented Germanium crystals. Each interaction of a γ-ray will result in a typical signal pattern on the segments surrounding the interaction point. The goal of pulseshape analysis (PSA) is to reconstruct not only the position but also the number of interactions and each energy deposit from the recorded signals. Because of the amount of data generated by the detector it is imperative that this analysis is carried out in real-time.

1. Machine-learning Algorithms

As stated by the *No-Free-Lunch* theorem there cannot be any algorithm that is optimal for every possible scenario [4]. This led to the idea to use Machine-learning algorithms, e.g. neural nets to classify the data into different categories and then use a specialized algorithm for each case. A reasonable choice for the different categories is the number of interactions that occurred. We look here specifically at the situation for neural networks(NN).

1.1 Neural Networks

NN's have one Input and one Output layer with a given number of hidden layers in between(Fig. 1). Each layer has N_k neurons holding a single value y_i given by
$y_i = g(\, b + \sum_{j=1}^{N_k} w_{ij} \cdot x_j \,)$.
Where w_{ij} is the weight for the connection from neuron j to neuron i, x_j the value from neuron j, b is a possible bias and $g()$ is the activation function. In a feed-forward NN layer i only has connection to neurons in layer $i-1$ [2].

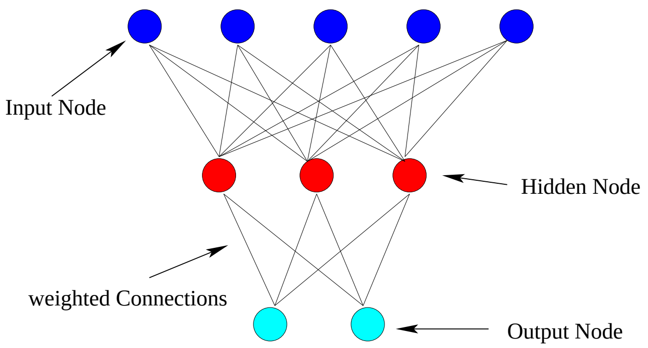

Fig. 1: Schematic design of a feed-forward neural network

A major concern when deciding what number of hidden layers and neurons to use is that the net may overfit the data. The manner in which this decision is made needs to be logically consistent and must take *Occam's Razor* into account. This can only be achieved through Bayesian Model Selection [1].

1.2 Training a Neural Network

A NN can be trained with data where the result is known. A common technique is *backpropagation*. For a given set of parameters the error of the nets output is assessed through a simple likelihood or a full bayesian posterior probability distribution. This error is then backpropagated to the other layers and the parameters are adjusted based on some learning rule [2,3].
For PSA we would get the probability mass function for the number of interactions as output from the net. If neuron k gives the probability for k interactions then for training data the values are 1 for the correct number and 0 else.

2. Feasibility for PSA

There are two general scenarios to consider. The first one, where a single net is used for the full crystal and the second one where one net is used per segment.

2.1 Full Crystal

Here the NN needs to be trained with data where the interactions occurred in any combination of segments. Multiple hits can happen in more than one segment. The number of possible combinations scales binomially with the number of interactions. For 5040 grid points (Ø 140 per segment, ~ 5 mm resolution) and 3 interactions there are $2.13 \cdot 10^{10}$ possible combinations. Adaptive techniques can reduce the data needed by using the probability for the combinations. However the different energy deposits add another layer of complexity. The amount of training data to be generated still easily combines to hundreds of TB.

2.2 Single Segment

In this scenario only a single segment was hit. The number possible combinations of points reduces to 447580 for 3 interactions. The signals scale linearly with energy so it suffice to consider only fractions of the total energy deposit. Consequently the above number needs to be multiplied with the possible combinations of energy fractions. To assess the probability for the different combinations of fractions multi-dimensional integrals have to be evaluated.

3. Conclusions

The use of classification algorithms for the full crystal is clearly not feasible. If we neglect the anisotropy of the crystal one NN could give sufficiently accurate results for the hits in a single segment. The accuracy will be worse for the asymmetric crystals. While NN's have been successfully implemented for single hits the computing time that would be necessary for training differently designed NN's for multiple hits renders the problem unfeasible.

References

[1] U.v. Toussaint *et al.*, AIP Conference, Melville, NY **735** (2004) 328-335
[2] C.M. Bishop,Clarendon-Press, Oxford, 1995.
[3] R.M. Neal,Technical Report CRG-TR-92-1, Departement of Computer Science, University of Toronto, 1992
[4] D.H. Wolpert,W.G. Macready, Technical Report SFI-TR-95-02-010, Santa Fe Institute, 1995