Chapter 4

Prediction of Location of High Energy Shower Cores using Artificial Neural Networks

4.1 Introduction

Several works exist which have used different approaches to analyze EASs and thereby develop applications suitable for EAS shower size prediction and location as referred to in Chapter 1. This work discusses the possibilities of using ANNs for individual EAS data evaluation. A work as cited in [Sommers 2004] uses ANN for providing a mass likelihood distribution for each measured shower, based on its multi-parameter training with simulated showers. Another work by A. Chilingarian et al [A. Chilingarian 2007] neural network models to recognize the experimental EAS without known primary energy.

The present chapter discusses the formulation and working of an ANN based system for prediction of EAS core location in the range $10^{10.5}$ to $10^{20.5}$ eV. The system uses twenty MLP - a class of feed forward ANN trained with error back - propagation algorithm to determine shower core coordinates. The twenty MLP cluster is trained in a cooperative configuration to provide optimized results. The set up is trained extensively and tested with data samples simulated resembling experimental conditions.

4.2 Basic considerations of the ANN

Artificial Neural Network (ANN)s are non-parametric prediction tools that can be used for a host of pattern classification problems [Haykin 2003] including face recognition. The application of the ANN considers two aspects. A MLP is constituted for this work with one hidden layer and
input and output layers. The choice of the length of the hidden layers have been fixed by not following any definite reasoning but by using trial and error method. For this case several sizes of the hidden layer have been considered. Table 3.1 shows the performance obtained during training by varying the size of the hidden layer. The selection of the activation functions of the input, hidden and output layers plays an important part in the performance of the system. The considerations of selection of activation fictional can be summarized by the Table 3.2.

The outcome of the MLP blocks vary depending upon the number of training sessions and the data used. Mean Square Error (MSE) convergence and prediction precision are used to ascertain the performance of the MLP blocks. Training of an MLP using (error) Back-Propagation is described in 2.

4.3 Application of ANN for Core Location Prediction

Showers were generated according to a modified NKG function [Hanna 1991] with particle shower primary energy in the range $10^{10.5}$ to $10^{20.5}$ eV with Moliere radius of 70 m. Their cores were evenly distributed within a circle of radius 50 m centered on the middle of the array and the detectors distributed within a circle of 100 meter radius. This restriction was adopted to avoid edge effect. A conceptual model of the core and detector locations used for the work are depicted in Figure 4.1. The high energy showers between $10^{10.5}$ to $10^{20.5}$ eV are simulated and density values calculated. While calculating density values core positions and locations of detectors are important. These coordinates are used to simulate the density values. The work considers twenty shower events taking place within a radius of 50 meters. The ANN is designed to accept density values obtained for twenty showers and provide coordinates of the EAS events of which the measurements are made. The density values captured by the detectors and coordinates of the shower events are unique hence require separate ANN predictors for each of the positions. For twenty shower events similar number of ANN are formed and a lay-out akin to the committee machine [Haykin 2003] is formulated. The set-up is essential for accurate prediction and optimization of the results. The training is carried out in a cooperative environment as to minimize the predicted error. The fundamental considerations gov-
Figure 4.1: Conceptual set-up used for simulation of particle density of EAS

Figure 4.2: Experimental Set-up
Chapter 4. Prediction of Location of High Energy Shower Cores

62 using Artificial Neural Networks

erning the working and parameter selection of the cooperative ANNs or committee machines can be explained using the following analysis [Rojas 1996] [Haykin 2003]:

Let a training set of m input - output pairs be \((x^1, t^1), (x^2, t^2), \ldots (x^m, t_m)\) be given and N networks are trained using this set of data. For simplicity, let for n-dimensional input there be a single output. Let for network functions \(f_i\) for a number of networks represented by indices \(i = 1, 2, \ldots N\), the cooperative or committee network formed generates as output given as

\[
f = \frac{1}{N} \sum_{i=1}^{N} f_i
\]  

(4.1)

The rationale behind the use of the averaging in the output of the cooperative or committee network as given by eq. 5.15 is the fact that if one of the constituent networks in the ensemble is biased to some part of the input samples, the ensemble average can scale down the prediction error considerably [Rojas 1996]. A quadratic error function can be computed from each of the error vectors \(e_i\) using the ensemble function \(f\) as

\[
Q = \sum_{i=1}^{m} [t_i - \frac{1}{N} \sum_{i=1}^{N} f_i]^2
\]  

(4.2)

Using matrix notation, the quadratic error can be expressed as

\[
Q = \left| \frac{1}{N} (1, 1, \ldots)E \right|^2 = \frac{1}{N^2} (1, 1, \ldots)EE^T (1, 1, \ldots) \quad (4.3)
\]

\(EE^T\) is the correlation matrix representing the error residuals. If each function approximation produces uncorrelated error vectors, the matrix \(EE^T\) is diagonal and the \(i^{th}\) diagonal element \(Q_i\) is the sum of quadratic deviations for each functional approximation, i.e \(Q_i = \|e_i\|^2\). Thus,

\[
Q = \frac{1}{N} \left( \frac{1}{N} (Q_1 + Q_2 + \ldots + Q_N) \right)
\]  

(4.4)

It implies that the total quadratic error of the ensemble is less by a factor \(\frac{1}{N}\) than the average of the quadratic errors of the total computed approximations. This holds only if \(N\) is not very large. If the quadratic errors are not uncorrelated, i.e if \(EE^T\) is not symmetric, a weighted combination of \(N\) functions \(f_i\) can be approximated as

\[
f = \sum_{i=1}^{N} w_i f_i
\]  

(4.5)
4.4 Experiential Results and Discussion

The weights $w_i$ must be computed in such a way as to minimize the expected quadratic deviation of the function $f$ for the given training set. With the constraint $w_1 + \Delta \Delta \Delta + w_N = 1$, eq. 4.3 transforms to

$$Q = \frac{1}{N^2}(w_1, w_2, ..., w_N)EE^T(w_1, w_2, ..., w_N)^T$$ (4.6)

Differentiating the above eq. 4.6 w. r. t $w_1, w_2, ..., w_N$, and using a Lagrangian multiplier $\lambda$ for the constraint $w_1 + \Delta \Delta \Delta + w_N = 1$, the above functional modifies to

$$Q' = \frac{1}{N^2}wEE^T + \lambda(1, 1, ...., 1)w^T$$ (4.7)

$$= \frac{1}{N^2}wEE^T + \lambda 1w^T$$ (4.8)

where 1 is a row vector with all its N components equal to 1. Setting the partial derivative of $Q'$ with respect to $w$ to zero this leads to

$$\frac{1}{N^2}wEE^T + \lambda 1 = 0$$ (4.9)

With simplification,

$$\lambda = \frac{1}{N^21(EE^T)^{-1}1^T}$$ (4.10)

The optimal weight set can be calculated as

$$w = \frac{1(EE^T)^{-1}}{1(EE^T)^{-1}1^T}$$ (4.11)

assuming that the denominator does not vanish. This method, however, is dependent on the constraint that $EE^T$ is not ill-conditioned.

4.4 Experiential Results and Discussion

Each of the twenty units of the ANN cluster is formed by cascade feed-forward networks - a variation of the MLP trained with back-propagation. The average data size for each of the block is fifty sets of $20 \times 100$ where 20 represents the number of shower cores and 100 denotes the number of density values recorded by the detectors. Noise between -3 dB and 3 dB are mixed to make the ANN cluster robust enough to deal with variations found from experiential works.
Figure 4.3: Location of shower events with detector and core positions generated by the ANN set-up for one set of events

Figure 4.3 shows the location of shower events as predicted by the ANN blocks with detector positions. While generating the coordinates for core events, the detector positions inside and outside the 50 meter circle are considered. The density values for a shower event captured are related to the coordinates of the detectors. The shower event, therefore, confined to any random position can be related to coordinates using density values from detectors. Following these considerations, four shower events are shown in Figure 4.3. It shows ANN estimated shower positions with expected theoretical locations. Another set of results (Figure 4.4) are generated for shower core positions using committee machines which show better success rates compared to unitary ANN blocks. The details of experiments performed using committee machines are explained in Chapter 5. These plots are shown in Figure 4.4. At the beginning of the training, the ANN arrangement is provided with density values of core events taking pale inside the 50 meter circle. After training up to 5000 sessions, the ANN blocks provide a shower core clustering of 20 events some of which fall outside the 50 meter circle. The success rate improves with increase in number of training sessions and at the end of around 12,000 iterations, the shower events converge within the 50 meter circle.
4.4. Experiential Results and Discussion

Figure 4.4: Location of shower events with detector and core positions generated by the ANN set-up for another event generated using the committee machine architecture as required. It shows that the ANN set-up with increase in training sessions, learns the patterns of grouping of shower events concentrated within the circle of 50 meter radius. The plot is for one event of which the density values are fed to the trained ANN set-up. After training with fifty sets of data the plotted values are generated as the average of twenty sets of inputs of which half are with noise variation in the mentioned range. The results show a success rate of around 95%. The above is repeated for another event and a similar success rate is obtained. A set of results are shown in Table 4.1 summarizing the training and testing processing. The location of all the twenty showers has also been generated using the ANN cluster as a whole. Initially as the training is limited to a few thousand session, the event cluster is spread inside and outside the 50 meter radius. The expected results are a grouping inside the fifty meter circle. As training sessions are increased with more number of samples, the predicted results start to cluster inside the 50 meter circle. Figure 4.5 shows a grouping generated by the ANN - cluster after 5000 sessions of training. The grouping clearly shows the location of shower events generated using density values placed inside the circle. A better
Table 4.1: Average performance derived during training of the ANN cluster for one shower event.

<table>
<thead>
<tr>
<th>SL Num</th>
<th>Epochs</th>
<th>Success rate in %</th>
<th>% Processing Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5000</td>
<td>94.1</td>
<td>20.8</td>
</tr>
<tr>
<td>2</td>
<td>10000</td>
<td>95.3</td>
<td>186.3</td>
</tr>
<tr>
<td>3</td>
<td>15000</td>
<td>96.2</td>
<td>318.2</td>
</tr>
<tr>
<td>4</td>
<td>20000</td>
<td>96.3</td>
<td>473.5</td>
</tr>
</tbody>
</table>

Clustering of the events recorded after 10000 iterations is shown by Figure 4.6. The number of training sessions have been extended to 20000 also but the best results are obtained around the 10,000 to 12,000 mark. Hence, testing results are derived from the ANN cluster trained up to this limit. A unitary ANN block instead of a cluster can also be sued for the purpose but prediction results are atleast 5% below than that generated by the cluster. Moreover, the cluster with its optimization capacity provides the best result out of a sample set of twenty applied to it. This is an advantage provided by the cluster. Moreover, as the shower core has its own unique location and density values, a unitary block shows best discrimination capacity only when applied for a single shower core location prediction. A plot of the success rates of the core location prediction is depicted by Figure 4.7. The training doesn’t improve much after 12000 sessions but training time increases.

### 4.5 Conclusion

The work is an attempt to use ANN based techniques to predict core locations. The experiential work carried out with a ANN cluster is found to be better suited for handling core location prediction and optimization. Shower event data are simulated and the extracted data are used for ANN training which is tested further to verify the extent of training. The system thus developed is a readily available tool which can provide nearly precise location details of showers of primary energy in the range \(10^{10.5}\) to \(10^{20.5}\) eV. The system has the potential to become a part of an experiential set-up for which it needs to be extended and modified further for real time applications.
4.5. Conclusion

Figure 4.5: Shower events of four cases predicted by ANN after 5000 sessions taking density values from 100 detectors

Figure 4.6: Shower events of four cases predicted by ANN after 10000 sessions taking density values from 100 detectors
Figure 4.7: Success rates generated by the ANN cluster predictor during 1000 to 20000 training sessions
Figure 4.8: $Kai^2$ distribution generated by the ANN cluster predictor during 1000 to 20000 training sessions