The Possibility of Using Artificial Neural Networks for the Estimation of Mass Composition of High-Energy Primary Cosmic Ray

Jan Malinowski*, **Grzegorz Gustek**†

*Department of High Energy Astrophysics, University of Lodz, ul. Pomorska 149/153, 90-236 Lodz, Poland
†Faculty of Physics and Applied Informatics, University of Lodz, ul. Pomorska 149/153, 90-236 Lodz, Poland
e-mail: jmalinowski@uni.lodz.pl

***Abstract*. This paper shows that artificial neural networks (ANN) ANN can be used for determining the type of particles of high-energy primary cosmic ray (i.e. its mass composition) initiating the EAS. The approach implemented here can be used, e.g., in the Auger experiment.**

We describe the details of the ANN construction and demonstrate that the program is correct and can be further used to solve physical problems. The network was taught and tested based on the data for the maximum of the EAS development (X_{max}) and primary energy of a particle initiating this EAS ($lg(E_0)$). The identification of particles based on X_{max} and $lg(E_0)$ resulted in around 80% of correct answers for the light mass composition and 99% for the heavy one. We have correct answer for the mass composition with domination of one type of particles, i.e. light or heavy. Otherwise, an additional parameters should be included as ANN input data.

***Keywords*: Artificial neural networks, ANN, Mass composition;**

I. The Objective of this Work

Registering multiple outcomes of interactions there is a problem how to identify the primary particle's type or its other parameter. Since the distributions of registered values are in some areas the same for two types of particles, algorithmic methods cannot be used in this case.

Artificial neural networks (ANN) have the ability of finding, matching and generalizing similar characteristics. This is why they are sometimes used in physics (see for example [1-3]). However, in a lot of problems where this tool could be very useful research does not take advantage of it.

In recent years, experiments on EAS initiated by high energy cosmic ray (CR) particles (of around $10^{18} - 10^{20} eV$) have been conducted. Very known examples are experiment Auger and little older ones, AGASA and Fly's Eye. One of their key objectives is to identify the type of primary particles coming from outer space. In these experiments, detectors register values which have various distributions. For different types of particles, there are areas where their distributions overlap.

This paper presents how ANN can be used to analyze data registered in detectors and determine the type of primary particles responsible for initiating the registered EAS. Former papers on using ANN to solving physical problems cited above regard particles registered in different energy region and regarding other research areas. Moreover, in the paper [4] described possibility using of ANN in the Auger experiment is mentioned but the paper does not report the whole analysis.

II. The ANN Construction and Testing

The ANN has neurons located in layers. Each neuron has the activation function. The shape of this function is conclusive for the output vale, entire (0 or 1) or floating point number [5, 6]. For solving our problem we used a nonlinear sigmoid activation function of the following form: $y = 1/(1+exp(-\beta j)$. The slope coefficient of the sigmoid, β, is important here since it allows to control the shape of the curve. The range of sigmoid arguments where the values of the function are from 0 to 1 is the most important during the teaching of the network.

To teach the ANN, we used the supervised teaching method, i.e. "with the teacher", and the backward error propagation method. The algorithm of backward error propagation [6] needs a teaching vector on the input and known output data. Teaching means the correction of weights such that errors made by single neurons, and thus by the whole network, decrease in consecutive steps until reaching a satisfactory precision level. The tests of the network were conducted on typical examples like teaching of the AND and XOR functions, identifying 26 letters and reproducing a sinusoid. A satisfactory result of these tests confirmed good work our implementation.

III. The ANN in Experiment

The network has to learn the identification of particles initiating the EAS based on the values of X_{max} received from the simulations. The simulation of the EAS is conducted using the CORSIKA program [7].

Our network will detects two types of particles with extreme values of mass number, protons - light particles and iron nuclei - heavy particles. The network is first taught based on the simulation data and then tested based on computation data. The neural network used in this task gives two values, i.e. there are two neurons with

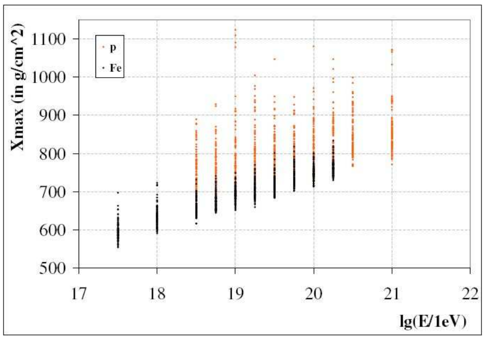

Fig. 1: The values of parameter X_{max} for the EAS depending on the logarithm of the primary particle enery, i.e. the particle initiating the EAS.

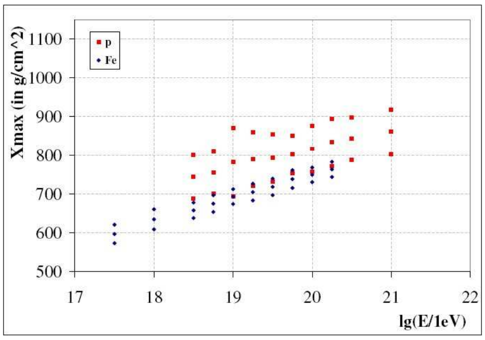

Fig. 2: Points of the teaching sequence $< X_{max} >$-σ, $< X_{max} >$, $< X_{max} >$+σ.

sigmoid function in the output layer. Each of them is supposed to identify one of two types of particles: proton or iron nucleus. Due to the sigmoid activation function, it is possible to identify the similarity of a primary particle to a proton or an iron nucleus.

The data includes X_{max} (divided by 10), i.e. the depth in the atmosphere where the shower has the maximum number of particles, and $lg(E_0)$, where E_0 is the energy of the particle initiating the EAS. X_{max} is divided by 10 because its values are much higher that the values of $lg(E_0)$ and the values given at the input should be of a similar dimension so that too large values alone do not decide about information processing in the network.

The data used for teaching the network is presented in Figure 1. This plot includes 20 different cases of simulation, 10 for protons and 10 for iron nuclei, each of them for a different energy. For each case 100 values of X_{max} have been given. Moreover, Figure 1 shows that a lot of points overlap what means that for similar data the network would be taught once that it is a proton and once that it is an iron nucleus. Thus the weights of nodes could eliminate one another during sequential iterations of the teaching process. To eliminate a large number of points to learn and still depict an interval of X_{max} values correctly, we estimated the average ($< X_{max} >$) and standard deviation (σ) of X_{max} for each of 20 data groups. These data describe an approximative interval of the values of X_{max} presented in Figure 2.

These approximate values are a sequence for the teaching process. Comparing them to the initial values in Figure 1, we can observe that they depict the intervals of X_{max} correctly. There are 60 teaching patterns, 30 for protons and 30 for iron nuclei. Although some patterns still overlap (similar to Figure 1), the number of them is very low and they occur in an area where the network should indeed have problems with the identification of particles. To indicate the significance of the value $< X_{max} >$ with respect to $< X_{max} >$-σ and $< X_{max} >$+σ, the value $< X_{max} >$ was given with different weights during the teaching. The results do not depend on the weights of $< X_{max} >$.

IV. Testing the Network

Testing of the network consists in drawing a certain number of values X_{max}, separately for protons and iron nuclei, for each $lg(E_0)$ from a normal distribution with a known and previously set mean of teaching vectors $< X_{max} >$ and a known standard deviation σ and checking how the network identifies the chosen values. In most cases, the network should identify the particle correctly.

Various configurations of the ANN have been checked, among others: 1-2, 2-2, 5-2 and 20-2 (the number of neurons in the first and second layer). The following ANN parameters have been analyzed: learning coefficient 0.15 or 0.1, momentum coefficient 0.4 or 0.3, weights and thresholds of neurons random from the interval (-0.1, 0.1) and the slope of sigmoid coefficient 0.1. The learning of the network was ended when the error decreased to a given value or the desired number of epochs was exceeded.

For an ANN with a 1-2 construction, the error after 20000 epochs amounted to 3.87. Figure 3a (the figure depicts only 10000 epochs) shows that the learning first evolved unsteadily and there were large changes of errors. After around 4000 epochs The error stabilized at the level of about 3.88 and then declined slightly. For an ANN with a 2-2 construction, the error after 20000 epochs was 3.87, i.e. its value was similar to the 1-2 network. Figure 3b presents the evolution of the learning process of this network. At first, the process is quite fast but after around 3500 epochs the error starts to fluctuate. The reason for such a result can be too large teaching and momentum parameters, too small number of neurons and connections among them and the fact that the network is taught to give different answers for similar values. Such cases are visible in Figure 2 as the points settled in a similar area but corresponding to different particles. For a 5-2 network, the teaching ended after 20000 epochs with an error of 3.64 equal for all patterns. This value is slightly better than error values in former examples. Figure 3c shows that the process of learning evolved quite steadily. It seems that this number of neurons is sufficient for teaching of the patterns presented above. The 20-2 network gave an error of 3.62 after 20000 epochs. The evolution of the teaching process was similar to the 5-2 network. As

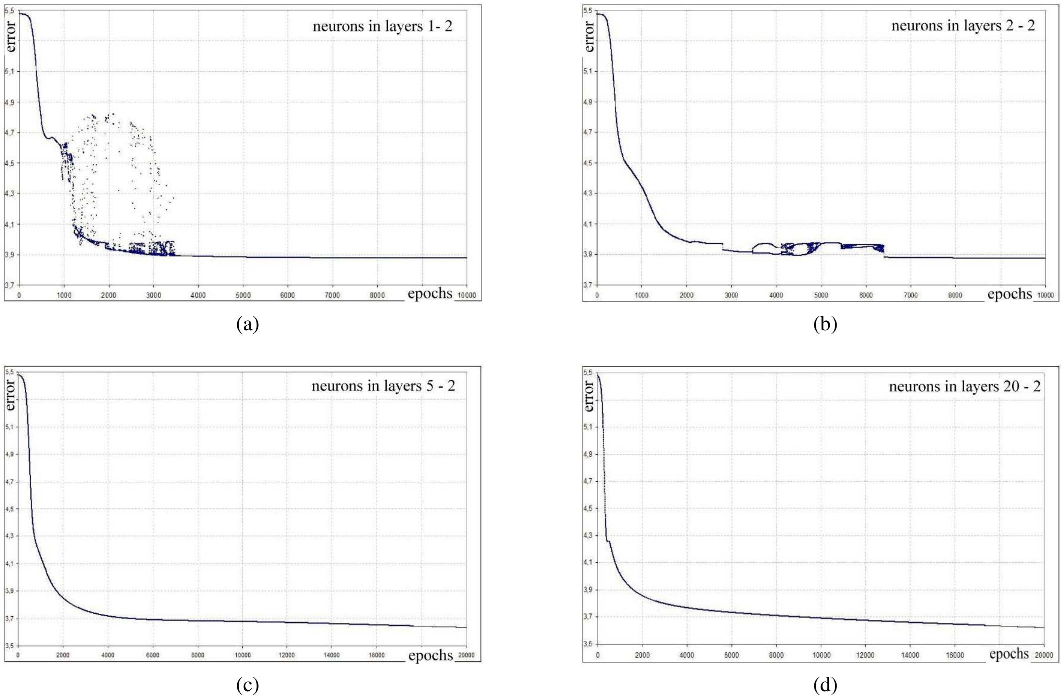

Fig. 3: Error depending on the number of epochs in the following ANN configurations (from above): 1-2, 2-2, 5-2 and 20-2.

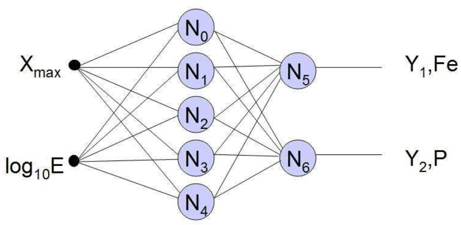

Fig. 4: Scheme of a 5-2 ANN used for identifying particles which initiate the EAS.

shown in Figure 3d a substantially higher number of neurons did not result in a large decrease of the error made by the network.

Consequently, achieving a much higher precision during the learning process is impossible for such patterns. The main reason is that they are quite similar for different particles. Too many connections between neurons may sometimes disturb the identification process. Moreover, too many neurons can cause that the network learns patterns very precisely and looses the ability of generalization. For these reasons and due to the testing results, we choose the 5-2 network for the further analysis. Figure 4 presents its scheme.

V. The Identification of Single Particles

In the second stage of tests, we establish the intervals for the values of outcoming signals for which a particle was classified as light (p) or heavy (Fe). The test involves pooling 1000 X_{max} for each $lg(E_0)$, as described in section 3. Further, the network's answers from the interval (0, 1) are put to a histogram. After the analysis of histograms, the identification threshold was chosen at 0.7. The cases with values from 0.3 to 0.7 are not identifiable, which corresponds to the situation in Figure 2, where the values of protons and iron nuclei overlap. The values lower than 0.3 mean that the second neuron returns the values higher than 0.7, i.e. the particle is identified incorrectly.

Table 1, columns 2 and 3, presents the result of identification for iron nuclei. It shows that for the threshold equal to 0.7, the identification by this network is very good for some energies, especially lower ones. For higher energies, the number of correctly identified particles decreases to around 60%. However, taking former observations into account, i.e. that the value interval of X_{max} for both types of particles overlaps, we can expect such identification efficiency. Table 1, columns 4 and 5, presents the result of identification for protons. It can be seen that the results of tests for protons are slightly worse. The identification rate for all energies is at the level of 60%-70%, which is an acceptable result. However, 15%-20% of incorrect identification is not satisfactory. The reason is, as in the case of iron nuclei, the same data values for some protons and iron nuclei (see the Figures 1 and 2). Such performance of the network can be thus expected.

VI. The Identification of Particle Groups

For teaching the network, the same data as in the last case was used. ANN has three inputs The following values are given to them:

TABLE I: The identification of primary particles type (in %) for iron nuclei (col. 2 and 3) and for protons (col. 4 and 5); each sample includes 1000 nuclei.

$lg(E_0/1eV)$	identified	not identified	identified	not identified
1	2	3	4	5
17.50	99.4	0.6	-	-
18.00	96.0	4.0	-	-
18.50	97.6	2.4	56.4	43.6
19.00	89.7	10.3	60.7	39.3
19.50	77.0	23.0	64.0	36.0
20.00	61.0	39.0	65.3	34.7
20.50	-	-	69.2	30.8
21.00	-	-	67.8	32.2

$lg(E_0)$ - logarithm of primary energy; $< X_{max} >$ - average value of X_{max} $S_{X max}$ - standard deviation.

Based on the experience from the exercises described above, we set the ANN architecture to 4-2: 4 neurons in the first layer and two neurons in the output layer, one for detection of the light composition and the second one for the heavy composition.

The network was taught with the parameters as below: u = 0.3 - learning coefficient; μ = 0.6 - momentum coefficient; β = 0.1 - sigmoid slope coefficient; (-0.25, 0.25) - starting weights, random from the interval; 0.05 - learning accuracy. The learning process was completed after 32597 epochs.

Different and extremal cases of the cosmic ray mass composition have been assumed for the analysis. Table 3 presents the percentage of nuclei for each case. The test was conducted for particles with energy $10^{19.5}$. Protons, He, N, Si and Fe nuclei have been taken into account. For the calculation of $< X_{max} >$ for these particles the following fit was taken:

$< X_{max} >= (1 - B)X_o(log(E/\epsilon) - log(A))$,

where: $X_0 = 37g/cm^2$, ϵ=81 MeV, B=0.47, A-particles mass number. The standard deviation of X_{max} distributions was fitted of σ for p and Fe nuclei simulation data. For all tests 1000 values of X_{max} were drawn from the normal distribution with the parameters presented in Table 2:

TABLE II: Simulation parameters.

Nuclei (A)	$< X_{max} >$	σ
p (1)	793	61
He (4)	765	47
N (14)	741	35
Si (28)	728	28
Fe (56)	714	21

TABLE III: Percentage of nuclei in mass composition; average and standard deviation of X_{max} for models.

Model	p	He	N	Si	Fe	$< X_m >$	σ
1	70.0	10.0	5.0	7.5	7.5	775.6	56.9
2	7.5	7.5	5.0	10.0	70.0	726.7	37.4
3	5.0	5.0	80.0	5.0	5.0	743.3	38.6
4	45.0	5.0	0.0	5.0	45.0	753.6	59.8
5	20.0	20.0	20.0	20.0	20.0	745.4	47.2

TABLE IV: The ANN answers for the assumed models (Ansv-Fe - identified by the ANN as heavy; Ansv-p - identified as light).

Model	Ansv-Fe	Ansv-p	Interpretation
1	0.009	0.991	Light
2	0.823	0.177	Heavy
3	0.489	0.511	not identified
4	0.010	0.990	Light
5	0.038	0.962	Light

Table 3 presents weighted averages $< X_{max} >$ and standard deviations σ for each model of mass composition. Table 4 shows that the ANN identified 3 of 5 cases correctly. Clearly, neural network works better when one type of particles dominates. In the case of equal number of light and heavy particles, ANN interprets them as protons. The reason is that the X_{max} distributions for protons have larger standard deviations than the ones for Fe. That is why protons are dominating in the case of partial overlapping of X_{max} distributions coming from p and Fe.

VII. Conclusion

The results shows that the ANN gives correct answers by the detection of the primary cosmic ray mass composition when the particles of one type, i.e. light or heavy, are dominating (82% - 99%). The correct answer, which is the lack of detection, is received also for the composition with the majority of nuclei from the Middle group, i.e. between p and Fe. Therefore, the ANN presented in this paper can be used when the primary cosmic ray includes a majority of only one type of particles, i.e. mass composition is heavy, light or middle. For the cases when the numbers of all types of particles in the mass composition is the same, the ANN (described in this paper) is not able of giving the correct answer. In such case, an additional parameter should be included as ANN input data. Such a parameter should characterize EAS as a function of the type of primary particle.

References

[1] R. Scrimaglio, *et al.* , *Analysis of sileye-3/Alteino data with a neural network technique: particle discrimination and energy reconstruction* Advances in Space Research, Vol.37, 2006, p.1697-1703

[2] S. Bussino, S. M. Mari, *Gamma-hadron discrimination in extensive air showers using a neural network* Astroparticle Physics, Vol.15, Issue 1, 2001, p.65-77

[3] M. Tamada, *Neural net classification of proton- and heavy nucleus-induced cosmic ray families* J. Phys. G: Nucl. Part. Phys., Vol.23, 1997, p.497-506

[4] M. G. Tanco, S. Sciutto, A. Tiba, *Neural Networks as a Statistic Diagnostic Tool for Mass Composition at the Highest Energies* Proc. of the 28th ICRC, Tsukuba, Japan, 2003, p.587

[5] L. Rutkowski, *Metody i techniki sztucznej inteligencji* PWN, Warszawa, 2006

[6] W. Duch, J. Korbicz, L. Rutkowski, R. Tadeusiewicz, *Sieci neuronowe* Exit, Warszawa, 2000

[7] D. Heck, *et al.* , FZKA 6019, Forschungszentrum Karlsruhe, 1998