MNRAS **000**, 1–7 (2019) Preprint May 22, 2019 Compiled using MNRAS LATEX style file v3.0

Optimizing neural network techniques in classifying *Fermi*-LAT γ-ray sources

M. Kovačević[1]★, G. Chiaro[2]†, S. Cutini[1], G.Tosti[3]

[1]*INFN–Istituto Nazionale di Fisica Nucleare Sez. Perugia, via Pascoli snc, I-06123 Perugia, Italy*
[2]*INAF–Istituto di Astrofisica Spaziale e Fisica Cosmica , I-20133, Milano, Italy*
[3]*Dipartimento di Fisica e Geologia, Univ. degli Studi di Perugia, via Pascoli snc, I-06123 Perugia, Italy*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Machine learning is an automatic technique that is revolutionizing the scientific research with innovative applications and the Artificial Neural Networks (ANN) is a powerful machine learning method widely use in astrophysics. In eight years of operation the Fermi-LAT gamma telescope detected more than 5000 γ-ray sources but the number of unassociated sources and uncertain blazars has exceeded 50% of the detected sources. ANN algorithms were applied to classify Fermi uncertain sources when strict classifications were not available significantly improving the number of classified objects. The aim of this study was to optimize the precision and effectiveness of an ANN machine learning method in order to open up new considerations on the population of the γ-ray sky, and the precision of significant samples selection planning for rigorous analyses and multiwavelength observational campaigns.

Key words: Methods: statistical – Galaxies: active – radiogalaxy objects: general – Gamma-rays: galaxies – radio continuum: galaxies

1 INTRODUCTION

Since August 2008 the Fermi Large Area Telescope (LAT) provides the most comprehensive view of the γ-ray sky in the 100 MeV - 300 GeV energy range (Atwood et al. 2009). The LAT 4-year Point Source Catalog *3FGL* (Acero et al. 2015) listed 3033 γ-ray sources of which 1717 were blazars considering 573 sources whose blazar classification remained uncertain. In addition 1010 of the detected sources had not even a tentative association with a likely γ-ray emitting source. As a total result, the nature of about half the γ-ray sources is still not completely known even if, because of blazars are the most numerous γ-ray source class, it could be reasonable to expect that a large fraction of unassociated sources might belong to BLLs or FSRQs. When rigorous classification analyzes are not available, machine learning techniques (MLT) represent powerful tools that enable identification of uncertain objects based on their expected classification. Machine learning is a data analytics technique that teaches computers to do what comes naturally to humans and animals: learn from experience. Traditional computer programs do not consider the output of their tasks and therefore they are unable to improve their efficiency. A machine learning method addresses this exact problems and involves the creation of an algorithm that is able to learn and therefore improve its performances by gathering more data and experience. MLT uses identified objects to teach the algorithm to distinguish each source class on the basis of parameters that describe its intrinsic features. The algorithm adaptively improves its performance as the number of samples available for learning increase. ANN algorithm has been applied to astrophysical objects, such as γ-ray unclassified sources, in order to quantify their probability of belonging to a specific source class. The algorithm generates an output that can be interpreted as a Bayesian a posteriori probability modeling the likelihood of membership class on the basis of input parameters (Gish 1990; Richard et al. 1991). In this work we explored the possibility to improve the performance of a machine learning algorithm based on the variability of blazars applying new physical parameters that characterize the nature of those sources and some statistical adjustments in order to increase the accuracy of the algorithm making it more efficient and effective. The expected result should be an optimized algorithm that is able to estimate, with more precision than in the past, the number of uncertain blazars that could belong to BLL or FSRQ class in the Fermi LAT Point Source Catalogs.

The paper is organized as follows: in Sect. 2 we provide a brief description of the main features of the most frequently used machine learning techniques in astrophysics. In Sect. 3 we present our optimization of an ANN method. In Sect. 4 we compare the performance of the optimized algorithm against the original one and comment the results testing the performances of the new algorithm sample

★ E-mail: milosh.kovacevic@gmail.com
† E-mail: chiaro@isaf-milano.inaf.it

of uncertain or unassociated γ-ray sources. We discuss predictions and implication of our results in Sect. 5.

2 MACHINE LEARNING TECHNIQUES

In previous studies (Ackermann, M. et al. 2012; Lee et al. 2012; Hassan et al. 2013; Doert et al. 2014; Chiaro et al. 2016; Mirabal et al. 2016; Saz Parkinson et al. 2016; Lefaucheur et al. 2017; Salvetti et al. 2017) and other authors have explored the application of machine learning methods classifying uncertain γ-ray sources. Here we briefly introduce the general features of the most frequently MLT used in astrophysics.

- **The Random Forest.** The Random Forest method (RF) (Beimann 1973) is an ensemble learning method that uses decision trees as building blocks for classification, regression and other tasks. By aggregating the predictions based on a large number of decision trees, Random Forest generally improves the overall predictive performance while reducing the natural tendency of standard decision trees to over-fit the training set.The Random-Forest package also computes the proximity measure, which, for each pair of elements (i, j), represents the fraction of trees in which elements i and j fall in the same terminal node. This can be used to calculate the *outlying-ness* of a source, as the reciprocal of the sum of squared proximities between that source and all other sources in the same class, normalized by subtracting the median and dividing by the median absolute deviation, within each class (Liaw et al. 2002). Saz Parkinson et al. (2016); Hassan et al. (2013); Mirabal et al. (2016) used the Random Forest algorithms classifying uncertain AGN or blazar-*like* sources from the Fermi gamma ray source catalogs (Acero et al. 2015; Ackermann, M. et al. 2015).
- **The Support Vector Machines.** The Support Vector Machine (SVM) (Cortes et al. 2013; Vapnik 1995) is a discriminative classifier formally defined by a separating hyperplane. In other words, given labeled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples. In two dimensional space this hyperplane is a line dividing a plane in two parts where in each class lay in either side. The method maximises the separation between different classes, which can then be used in classification or regression analysis. In Hassan et al. (2013) the authors used a Support Vector Machines algorithm with the Random Forest algorithm building a classifier that can distinguish between two AGN classes: BLL and FSRQ based on observed gamma-ray spectral properties.
- **The Boosted Decision Trees.** The Boosted Decision Tree (BDT) was used in Freud et al. (1999); Lefaucheur et al. (2017). It is based on the decision trees, a classifier structured on repeated yes/no decisions designed to separate *positive* and *negative* classes of events. Thereby, the phase space of the discriminant parameters is split into two different regions (Freud et al. 1999) and generates a forest of weak decision trees and combines them to provide a final strong decision. At each step, misclassified events are given an increasing weight. BDT was used in with another method of the artificial neural network class of algorithms to classify uncertain blazar sources among the Fermi-LAT detentions.
- **The Artificial Neural Network.**

The Artificial Neural Network (ANN) (Bishop 1995) is probably the most used machine learning technique in astrophysics. ANN algorithms were used in Chiaro et al. (2016); Salvetti et al. (2017) for classifying uncertain γ-ray sources and were also used in the above cited Saz Parkinson et al. (2016); Hassan et al. (2013); Lefaucheur et al. (2017).

Standard neural network consists of an input layer, one or more hidden layers and an output layer. In Fig. 1 the schematic view of the basic architecture of an Artificial Neural Network algorithm is shown. Neurons in the input layer are values of input parameters. Each neuron in the first hidden layer has a set of *weights* which are associated to input parameters. The sum of products between each weight and input parameter[1] is then used in a *transfer function* to create a single output. The outputs of neurons in hidden layer are then used as input values for the neurons in successive hidden layer and so on. Neurons in output layer work the same way and they give the final result.

When classification is the goal, usually the number of neurons in output layer equals the number of classes; the sum of outputs from these neurons equals 1 and output value from each neuron is interpreted as probability of belonging to a given class.

Training the network with known/labeled sources involves setting the weights of all neurons in the network so that difference between given outputs and desired outputs, quantified by a *Loss/Cost function*, is minimized. The sample of sources used in training the network typically contain 50 - 80 % of all known sources. The rest are divided into two independent samples - *validation sample* and *test sample* which are used to avoid over-fitting and to evaluate network on sources it hasn't seen during training.

The original ANN algorithm that we considered in this study has been used for the first time in Chiaro et al. (2016) (hereinafter *C16*) and subsequently in Salvetti et al. (2017) (hereinafter *S17*). The algorithm compared the γ-ray light curve of the source under investigation with a template of classified blazar class light curves, then measured the difference in a proper metric. The authors of both the paper used a simple neural model known as Two Layer Perceptron (2LP), rather similar to the method used by Lefaucheur et al. (2017) but with a simpler architecture.

In this work we explored possibilities to improve the efficiency of the original algorithm used in *C16* and subsequently in *S17*. Even if the original ANN algorithm was very effective, the number of sources with uncertain classification in *C16* and *S17* remained consistent. Optimizing the original algorithm we used *TensorFlow*[2] which is an open source library for machine learning. It provides details on the process of training and is built to be fast - network can be trained on a computer in relatively short time.

3 THE METHOD

3.1 Network optimization

In *C16* the Empirical Cumulative Distribution Function (ECDF) of the monthly bins of the 3FGL BCUs γ-ray light curves was applied to the ANN as an estimator able to classify BCUs into BLLs and FRSQs. This parameter contains information of flaring patterns, along with maximum flare, and variability of the sources. The distinctiveness of 3FGL BLLs and FSRQs sorted by ECDF is shown in Fig. 2. This result convinced the authors to use the ECDF as the sole ANN parameter to compute the likelihood of their sample of uncertain sources to be BLL or FSRQ. However in *C16* analyzing 573 BCUs, 77 sources remained of uncertain classification. Also in *S17* classifying with the same algorithm the AGN-*like* sources, 103 of 559 sources remained of uncertain blazar classification.

[1] A single value *bias* can be added.
[2] https://www.tensorflow.org

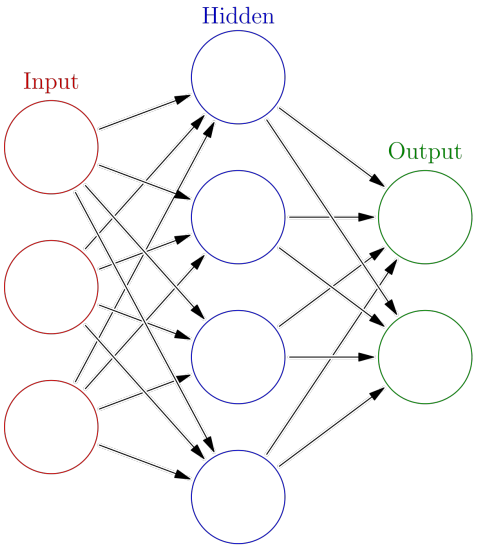

Figure 1. Schematic view of a Two Layer Perceptron (2LP), the Artificial Neural Network architecture. Data enter the 2LP through the nodes in the input layer. The information travels from left to right across the links and is processed in the nodes through an activation function. Each node in the output layer returns the likelihood of a source to be a specific class.

In this study, in order to improve the performance of the network, we considered and tested new and different parameters as potential input for the algorithm.

3.2 Flux in different energy bands

An interesting set of parameters that we considered was the time integrated fluxes in different energy bands Fig. 3. This set of parameters contains information of average spectral index, hardness ratio and peak energy. In Lefaucheur et al. (2017) it was highlighted as a good selector. We considered five different energy bands : 100-300 MeV, 300-1000 MeV, 1-3 GeV, 3-10 GeV, 10-100 GeV. and in Fig. 3 3FGL BLL and FSRQ distributions in energy bands against the time-integrated γ-ray flux (cm^{-2} s^{-1}) are shown. It is interesting to note that in the range of flux $\sim 10^{-10}$ and energy bands from 100-300 MeV up to 1-3 GeV mostly BLLs are present (see diagram below left in Fig. 3) while for energy band 10-100 GeV FSRQs are more numerous for the lower values of fluxes, around $\sim 10^{-15}$. It is also notable that majority of BLLs and FSRQs have different slopes across all energy bands Fig. 3, which is in part a reflection of different average power-law indices.

3.3 Radio and X-ray fluxes

Looking beyond γ-ray features of blazars, interesting information could be obtained from a multiwavelenght study of the sources and particularly from X-ray and radio flux. In this study we also tested the possibility to use those two parameters improving the performance of the network. We didn't consider any optical spectroscopy data because when considering uncertain sources, optical spectra are very often not available or not sufficiently descriptive of the nature of the source.
A particularly interesting parameter seems to be the ratio of the radio (Sr) flux versus the X-ray flux. In Fig. 4 the radio and X-ray photon flux histograms are shown. When both the parameters are

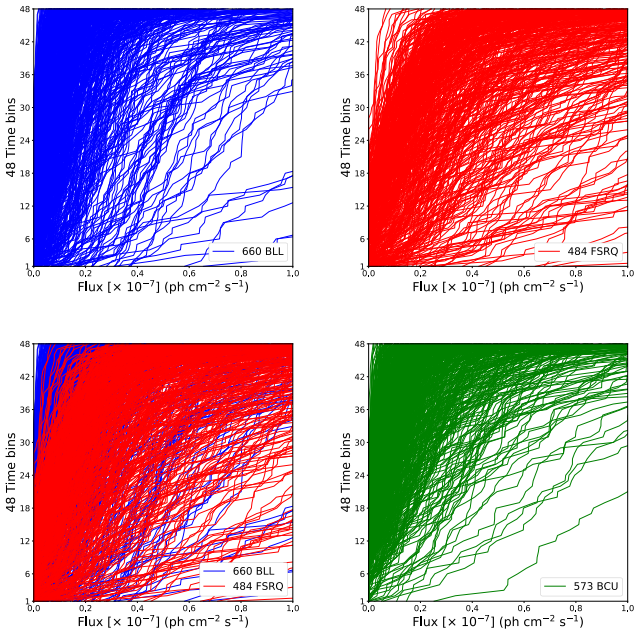

Figure 2. ECDF light curves with monthly time bins of 3FGL blazars from 4 years of data. Monthly flux values for each source are sorted from lowest to highest. Each curve represents a single source. BLLs are in the top-left, FSRQs in the top-right, both are in the lower-left and BCUs are in the lower-right. ECDFs for some sources extend beyond the plot limit of 10^{-7} cm^{-2} s^{-1}. *Credit* : (Chiaro et al. 2016)

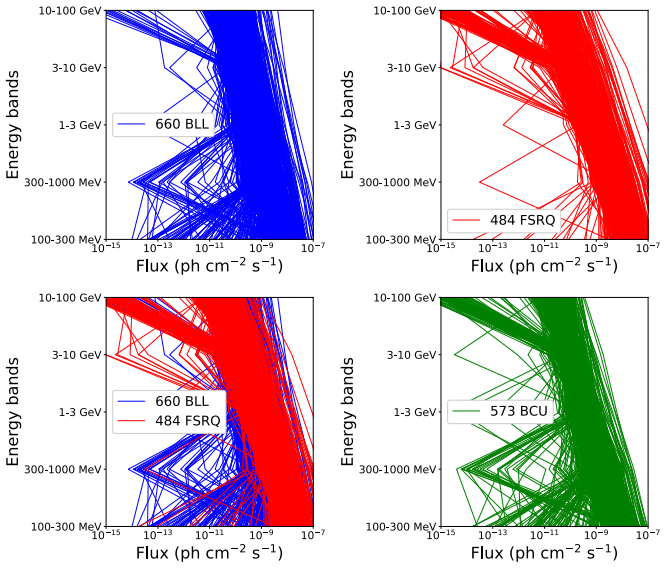

Figure 3. Time integrated fluxes in 5 different energy bands. Each curve represents a single source. BLLs (blue) are in the top-left, FSRQs (red) in the top-right, both are in the lower-left and BCUs (green) are in the lower-right.

considered separately the contamination is not negligible (histogram on top and in the middle), but when the ratio Sr/X-ray is considered it is possible to distinguish a clean area for BLLs by ratio values minor than 4×10^{13}. Unfortunately not all the known and uncertain blazars have both radio and X-ray flux data. However the final result is still appreciable because considering 3FGL blazars, 322 BLL out of 660 have both radio and X measurement and for 188 sources (28%) the value is minor than 4×10^{13}. Also 146 FSRQs out of 484 have both measurements but none shows a ratio value minor of the defined threshold. Finally, out of 507 3FGL BCUs, 107 sources have both measurements while 57 show a value of the ratio minor than the threshold. This means that the ratio of radio to X-ray flux, although an overlap of data in higher values in not negligible it could be considered as a *smoking gun* area where a good separation for BLLs is possible, if the ratio value is minor than 4×10^{13}.
Radio fluxes used were measured at frequencies of 1.4 GHz and 0.8 GHz; the X-ray fluxes were measured in the 0.1 - 2.4 keV range. Data were obtained from the third Fermi-LAT 4-year AGN Catalog *3LAC* (Ackermann, M. et al. 2015).

3.4 ECDF data

Since ECDF curves represent the main set of ANN parameters originally used in *C16*, it was interesting to test if some statistical methods could improve the final performance of the network. In Fig. 2 by a simple *looking by eyes* observation of the plots on the left side it is possible to see on the top-left of the diagram a specific area where the overlap between BL Lac and FSRQ is minimal. This area shows an upper limit of the γ-ray flux of ~2.5 x 10^{-8} cm^{-2} s^{-1} that leaded to a first *qualitative* recognition of BLL candidates. This result seems reasonable if we consider the different flare activity of the two class of blazars and the related different slope of the ECDF plots. However for sources with higher values of photon flux, the light curves of BLLs and FSRQs are much more intertwined and they don't seem to be easily distinguishable.
Removing some of these sources would help the network make a better separation. One way to proceed is to remove sources which don't have a single monthly time bin without flux value above detection threshold Fig. 5. Applying this constrain, number of BLLs fell from 660 to 589 (-10%) and FSRQs from 484 to 433 (-10%) . The reduction of source number did not affect training and testing the network. Also BCUs number fell from 573 to 567, so only 6 sources were lost for classification.

3.5 Data input

If certain parameter values (for an input neuron) from different sources vary over several orders of magnitude, it is usually good practice to use the logarithm of those values. However it is not always the case that this strategy improves the performance of the network since all input neurons are treated together in training of the network. Combinations with both original values and their logarithms were checked. The results are reported in Sect. 4.1.
The input data were normalized by subtracting their average value and dividing by their standard deviation so most of the input values fell between -1 and +1 for each input neuron.
Majority of sources don't have γ-flux value above detection threshold for all monthly time bins. Additionally the radio (Sr) and X-ray data are missing for many sources. One way to deal with missing input data to neural network is to set those inputs to zero. In this way, zero input acts as if there is no input neuron. Since Fermi satellite

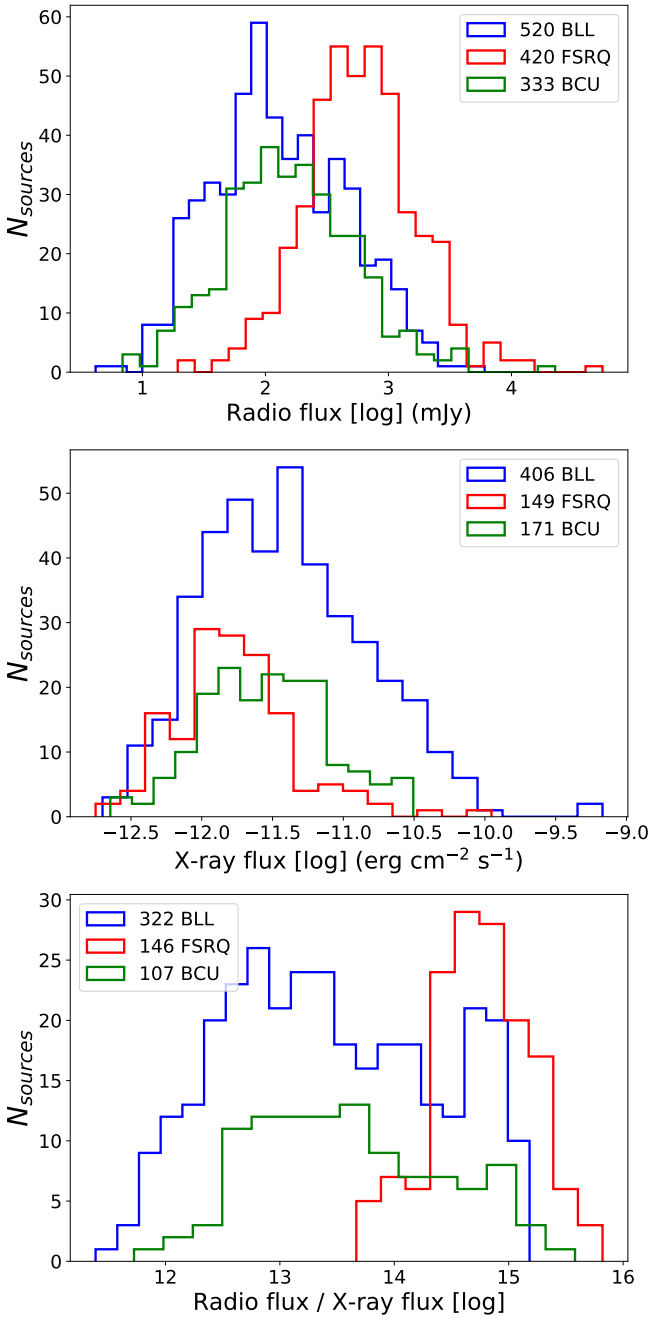

Figure 4. On top: 3FGL BLL and FSRQ radio flux histogram. Even if the two peaks are well separated, the contamination is not negligible. In the middle: 3FGL BLL and FSRQ X-ray flux histogram. The two classes show a full overlap and they are mainly indistinguishable using only this index. On the bottom: 3FGL BLL and FSRQ radio flow / X-ray flux ratio histogram. A BLL clean area is distinguished by ratio values lesser than 4×10^{13}.

sweeps the sky continuously, the non-detection of γ-flux is due to low value of photon flux during the detection and not observational constraints. The missing Sr and X data are due to low flux and/or observational constraints.
In our tests the input radio and X-ray data for sources with missing radio and X values were set to zero after normalization of input parameters (which they didn't affect). They were set to zero as well

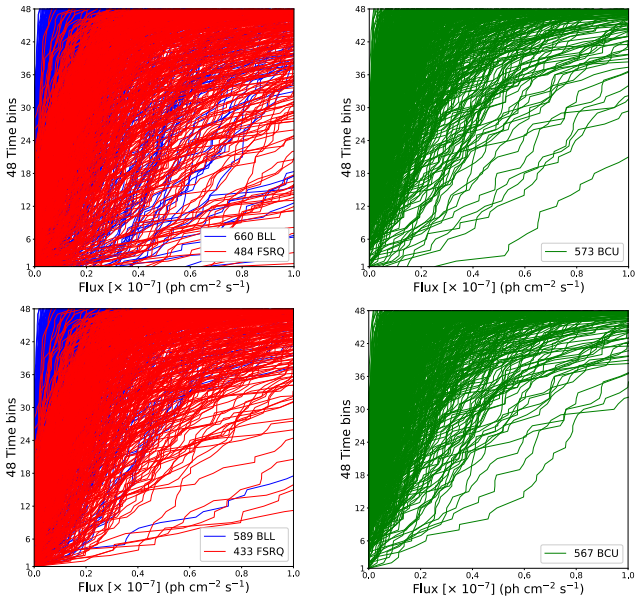

Figure 5. ECDFs: The upper two plots contain all the sources while the two plots at the bottom contain 3FGL blazars after applying the cut. The number of sources for each class is written on the plots. Blue curves correspond to BLLs, red to FSRQs and green to BCUs. After applying the cut, the lower-right part of the ECDF plot for BLLs and FSRQs becomes more clear. For BCUs, the lower-right part of the ECDF plot remains similar after the cut because there were not many sources passing trough it (only 6 sources are removed by the cut).

if the logarithm of Sr and X was taken.
The input data for missing monthly gamma-flux is zero but it wasn't set to zero after normalization (which it affected). After normalization these data had the lowest values. In case logarithm was taken, to avoid logarithm of 0, these data were set to minimal value of monthly detected flux of 4×10^{-11} cm^{-2} s^{-1} (of any source) before logarithm was taken. After that data were normalized as in the previous case.

3.6 Network architecture

What follows is the architecture of the optimized network used in this work.
The number of neurons in input layer is equal to the number of input parameters. The ECDFs accounted for 48, γ-fluxes in energy bands accounted for 5, ratio of radio to X-ray flux for 1 neuron. With all parameters used, the number of input neurons is 54.
The hidden layer consisted of 100 neurons. The choice was made by experimenting with single example for fixed number of training epochs. It was found that number of neurons should be higher then number of input parameters (about 50 in our case) but after that the performance didn't change noticeably with further increase. The output layer consisted of two neurons. The activation function for hidden layer neurons was hyperbolic tangent while for output neurons it was softmax which insured that the sum of output neurons equals 1.
The batch size was set to number of sources in training sample which insured smooth convergence and no fluctuation due to different sources in different batches. Although binary cross-entropy Loss function is commonly used for binary classification, mean square error produced better results in our case. The minimization algorithm used was adam-optimizer[3] (Kingma et al. 2014) with default values that converged quicker and gave better results in our case than classical stochastic gradient descent.

3.7 Training strategy

Typically samples/sources for training set and other sets are chosen randomly. The fluctuation in performance depending which sources are taken might be important. In our case there are about 1000 labeled sources (BLLs and FSRQs) and it was found that the number of unclassified BCUs may vary significantly depending how training and other sets are chosen. In Lefaucheur et al. (2017) the same problem was noted and we decided to test the strategy suggested from the authors, by training the network for 100 different training and testing samples, and then selecting the set which is closest to average results. The train set consisted of 70% and test set of 30% of the 3FGL classified blazars.
Aside from training network on 100 different train and test samples, to avoid introduction of second independent sample (with yet smaller number of sources in it), next strategy was used: the number of epochs was fixed for all combinations of input parameters and selections of training and testing samples, and network evaluated at the end; regularization was used to avoid over-fitting. The value for regularization was chosen so that it allowed network to get close to lowest test Loss function and to have it smoothly converge by the final epoch.
The desired outcome for training sample sources was set to [1,0] and [0,1] for BLLs and FSRQs respectively. In this way the output neurons returned the likelihood of source belonging to either class. Inputting parameters from known/labeled sources from testing sample into the trained network enables network evaluation. Two output neurons produce likelihood of source being BLL L_{BLL} or FSRQ L_{FSRQ} such that $L_{BLL} + L_{FSRQ} = 1$ for each source.
Network performance was evaluated by how many BCUs are left unclassified applying a 90% precision threshold *C16*.

4 VALIDATION

4.1 Different combination of input parameters

The set of parameters of our optimized algorithm are 48 γ-flux values from ECDFs Sect. 3.1, the 5 time integrated γ-flux values in 5 energy bands Sect. 3.2, the radio to X-ray flux ratio described in Sect. 3.3, and option of removing sources by applying a cut as described in Sect. 3.4. Applying those parameters we improved the performance of the original algorithm decreasing the number of unclassified BCUs. Considering the 3FGL blazars and BCUs and running the algorithm applying all the new parameters the ANN returned a number of unclassified BCUs of 30 instead of 77 as reported in *C16*. The application of the logarithm to any of the parameters instead worsens the network performance. When radio (Sr) and X-ray parameter were inputted separately instead of the Sr to X-ray ratio, the number further falls to 19. This result is due to the fact that more BCUs have separate Sr and/or X than both Sr and X. Out of the total 573 BCUs in 3FGL, 333 have radio measurements, 171 have X-ray measurements. Number of BCUs

[3] https://www.tensorflow.org/api_docs/python/tf/train/AdamOptimizer

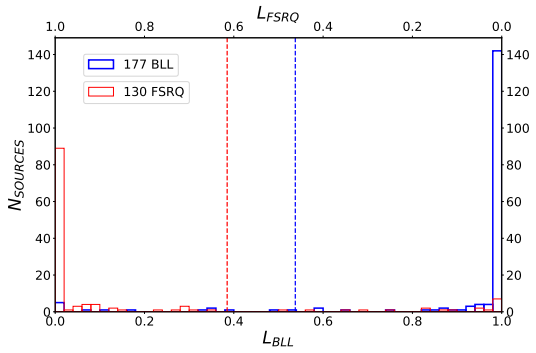

Figure 6. Histogram of L_{BLL} for BLLs and FSRQs from the testing sample. Number of BLLs in testing sample is 177 and FSRQs 130. The blue and red vertical lines (at $L_{BLL} = 0.542$ and $L_{BLL} = 0.389$ respectively) present thresholds for BLLs and FSRQs such that precision of 90% is obtained.

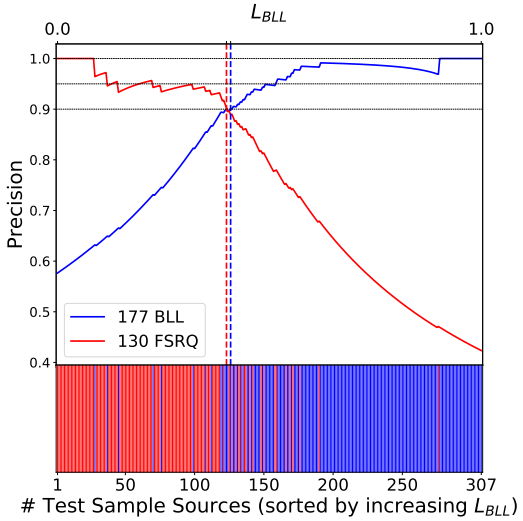

Figure 7. Lower bar present BLLs (blue) and FSRQs (red) sources from test sample sorted by increasing L_{BLL} and at equal distance from each other. The L_{BLL} doesn't increase linearly in the plot. The upper plot presents change of precision with L_{BLL} threshold for BLLs and FSRQs. The threshold where precision reaches 0.9 for BLLs and FSRQs is $L_{BLL} = 0.542$ and $L_{BLL} = 0.389$ respectively (blue and red vertical lines). Precision is on average monotonically increasing/decreasing function with L_{BLL} for BLL/FSRQ. The zig-zag oscillations in precision is due to finite and relatively small number of sources in testing sample.

which have both measurements is 107 while 397 have one or both. If a ratio is added as a third parameter to radio and X-ray alone, the performance remains the same.
Finally after the set of parameters was chosen, further tweaking the network, mainly by selecting mean absolute error as Loss function, helped by decreasing number of unclassified BCUs to 16.
All the number of unclassified BCUs mentioned above are average values of unclassified BCU from 100 different selections for training and testing samples.

4.2 Final results

In Fig. 6 histogram of L_{BLL} for BLLs and FSRQs from the optimized testing sample is presented. As expected BLLs concentrate towards $L_{BLL} \rightarrow 1$ while FSRQs $L_{BLL} \rightarrow 0$.

The precision of the optimized ANN considering a threshold of 0.9 can be seen in Fig. 7. Sources from test sample are sorted by their L_{BLL} (like in the Fig. 6), but sources are at equal distance from each other and L_{BLL} doesn't increase linearly. The threshold where precision reaches 0.9 for BLLs and FSRQs is $L_{BLL} = 0.542$ and $L_{BLL} = 0.389$ respectively (blue and red vertical lines).

Applying the threshold values of $L_{BLL} = 0.542$ and $L_{BLL} = 0.389$, as above described, the ANN leaves 16 BCU unclassified, 378 classified as BLLs and 173 as FSRQs.

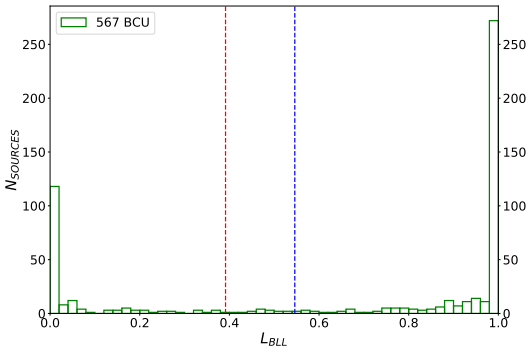

Figure 8. Applying trained network to 567 BCU sources. Applying the threshold values of $L_{BLL} = 0.542$ and $L_{BLL} = 0.389$ (blue and red vertical lines) leaves 16 BCU unclassified, 378 classified as BLLs and 174 as FSRQs.

5 CONCLUSION

In this study we explored the possibilities to increase the performance of a neural network method previously used for the classification of uncertain blazars. We considered the improvement of performance applying new parameters both from the physical features of the sources, and also from statistic adjustment of the algorithm input. We developed an optimized version of the original algorithm improving the selecting performance of 79%. The final result of this study left 16 uncertain blazar sources instead of 77 in Chiaro et al. (2016). This result confirms the machine learning techniques as powerful methods to classify uncertain astrophysics objects and particularly blazars. Given that the forthcoming *Fermi*-LAT catalog could show an increasing number of uncertain blazars. ANN technique could be a very interesting opportunity for the scientific community to select promising target for multiwavelenght rigorous classification and related studies at different energy ranges mainly at very high energies by the present generation of Cherenkov telescopes and the forthcoming CTA. [4]

6 ACKNOWLEDGMENTS

Support for science analysis during the operations phase is gratefully acknowledged from the Fermi-LAT collaboration for making the Fermi-LAT results available in such a useful form, the Institute of Space Astrophysics and Cosmic Physics in Milano -Italy (IASF), the Department of Physics and Geology of the University of Perugia

[4] The CTA project is available at `www.cta-observatory.org`

- Italy, National Institute for Astrophysics (INAF) Rome - Italy, and National Institute for Nuclear Physics (INFN) Rome - Italy.

References

Abdo, A. et al, 2010, ApJ, 720, 912-922.
Acero, M. et al., 2015, ApJS, 218, 2, 23 - 41.
Ackermann, M. et al., 2012, ApJ, 753, 83
Ackermann, M. et al., 2015, ApJ, 810, 14 - 34
Atwood, W. B. et al., 2009, ApJ, 697, 1071
Beimann, L., 1973, Statistics. With a view toward applications, Houghton Mifflin Co., Boston, 1973
Bishop, C. M., 1995, Neural Networks for Pattern Recognition
Bonning, E. W. et al., 2009, AAS, 41, 686
Chiaro, G. et al., 2016 MNRAS 462.3.3180C
Cortes, C. et al., 1995, Machine Learning, 273
Doert M. and Errando M., 2014, ApJ, 782, 41
Freud, Y. et al., 1999, 14, 771
Friedman, J. et al., 2000, AnSta, 28, 337
Gish, H., 1990, Proceeding on Acoustic Speech and Signal Processing, p. 1361
Hassan, T. et al., 2013, MNRAS, 428, 220
Kingma, D. P. and Ba, J., 2014, arXiv, 1412.6980
Lee, K. J. et al., 2012, MNRAS, 424, 2832
Lefaucheur, J. and Pita S., 2017, A&A, 602, 86
Liaw, A. and Wiener M., 2002, Classification and Regression by Random Forest. R News, 2, 18-22
Mirabal, N., et al., 2016, ApJ, 825, 69
Richard, M. P. and Lippman R. P., 1991, Neural Computation, 3, 461
Saz Parkinson, P. et al, 2016, ApJ, 820, 2
Salvetti, D. et al., 2017, MNRAS, 470, 1291
Vapnik V.,1995, The Nature of Statistical Learning Theory. Springer, New York, p. 138