Available online at http://www.journalijdr.com

IJDR

ISSN: 2230-9926

International Journal of Development Research
Vol. 10, Issue, 03, pp. 34645-34650, March, 2020
https://doi.org/10.37118/ijdr.18530.03.2020

RESEARCH ARTICLE

OPEN ACCESS

AUTOMATIC INCREMENT IN A KNOWLEDGE BASE BY MEANS OF FUZZY SYSTEM WITH SUPERVISED MACHINE LEARNING

*Ernande F. Melo, Muller, W., Manoel S.S. Azevedo and Hiram C. Amaral

University of the State of Amazonas

ARTICLE INFO

Article History:
Received 17th December, 2019
Received in revised form
21st January, 2020
Accepted 02nd February, 2020
Published online 31st March, 2020

Key Words:

Integration, Machine Learning, Fuzzy Systems, Knowledge Base, Classification.

****Corresponding author:* Ernande F. Melo**

ABSTRACT

The integration between Machine Learning (ML) and Fuzzy Systems is a recurring theme in the field of Artificial Intelligence (AI), specially regarding the deductive methods of a Fuzzy System, and, on the other hand, the inductive ones of ML. This article presents an experiment which integrates both approaches, thus showing that they may indeed be complementary. The experiments consists of providing a ML with a mechanism for automatic increment of its knowledge base by means of inserting examples (correctly classified by a Fuzzy Sistem) into supervised learning problems. Increment by means of inserting correctly classified examples allows for a growth of the base and an increase in the ML performance. Finally, in this experiment we show that (under certain conditions), a Fuzzy System ensures the correctness of those examples which will be inserted into the said base and thus ensures an increase in the ML performance.

Citation: ***Ernande F. Melo, Muller, W., Manoel S.S. Azevedo and Hiram C. Amaral,* 2020.** "Automatic increment in a knowledge base by means of fuzzy system with supervised machine learning*", International Journal of Development Research*, 10, (03), 34645-34650.

INTRODUCTION

ML applications are reaching a surprising number of fields, which in their turn, have been encouraging an increasing growth of research into ML. Extreme Machine Learning (EML) Huang, 2011; Huang, 2006; Zhai, 2016), Deep Learning Yann, 2015; Zhang, 2017), and its integration with Fuzzy Systems (Neuro-Fuzzy) BISWAS, 2016; Cpałka, 2013), are some examples of the developments in ML. In spite of these new approaches, the problem of classification (in supervised learning) essentially remains the same, that is, the learning method is inductive Hüllermeier, 2015). A ML learns if its performance increases with the growth of its knowledge base in relation to a class of tasks" Mitchell, 1997). The growth of the knowledge base, in its turn, occurs by means of its increment by means of the insertion of new correctly classified examples. The main problem is: how can we ensure that those new examples shall be correctly classified? Even though, it may appear to be a problem regarding estimating the degree of trust of reliability or uncertainty in a Deep Neural Network (DNN) Chollet, 2018; LeCun, 2015; Good fellow, 2016) (there being several ways to estimate reliability of a classifier, be it with black box, autoencoder Good fellow, 2016), or softmax white box Agarap, 2019)), this article presents the assurance of the correctness of the examples classified by a Fuzzy System (under certain conditions). In this context, some questions arise, the first one being related to the use of a Fuzzy System as a validator of classfications made by a ML. That is, if a Fuzzy System correctly evaluates the classification of an example, then the ML is not necessary. However, as we shall see, a Fuzzy System will not always act, but each time it is activated, the knowledge base will be incremented with the insertion of the correctly classified example by the Fuzzy System. Another problem is this: our article is not about Machine Learning or Fuzzy System, but rather about integrating two environments, exploiting the deductive methods of a Fuzzy System and the inductive methods of ML. In Section II we show two environments, a ML and a Fuzzy System. We briefly describe their basic components. In section III, we show the material and methods employed. Section IV, in its turns, presents the results of the experiment, Section V, Conclusions, and finally we present our acknowledgements and our thanks in Section VI.

Environments: In this section we shall briefly describe the basic components of a LM and of a Fuzzy System.

Machine Learning (ML): Figure 1 show the scheme of a ML, which consists basically of a knowledge base and a learning algorithm. Data are structured in attributes and (in the case of supervised learning) a class of each set of attributes Géron, 2017 and Kotisiantis, 2007). Part of such data is used for training the learning algorithm while another part is used for testing and evaluating the LM performance. The more data used for training it, the better is the LM performance and the more data in training may then be obtained by increasing the base.

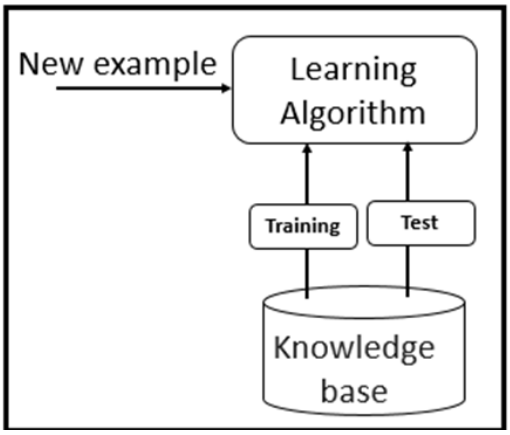


Figure 1. Schematic representation of a LM which is composed of a knowledge base with a set of data (for training) and another set for tests, as well as of a learning algorithm

Amongst learning algorithms, the most popular ones are: KNN, Tree, SVM and RNA Géron, Aurélien, 2017; Kotisiantis, 2007). For the present experiment, we chose the tree learning as our learning algorithm (but, for the purposes of this work, the algorithm selected is not relevant).

Fuzzy Systems: Figure 2 shows the scheme of a Fuzzy System Azeem, 2012; Zimmermann, 2010). The basic elements that make up the system are:

***Fuzzy Curves*:** Fuzzy Curves, described by relevance functions express the knowledge base on the problem. They are built in the process of fuzzifying, which converts quantitative variables into linguistic variables. A fuzzy curve relates quantitative variables to the degree of relevance in all linguistic variable curves. Figure 3 shows an example of flowers classification which shall be describe in section III of this article. The curves in the graph represent the degree of relevance of the attributes in each class. In the example given, the attribute is petal_length and the graph describe the degree of relevance of the attribute in each of the classes that were defined as linguistic variables (setosa, virginica and versicolor). The making of fuzzy curves or relevance functions usually requires an expert, who will definee the shape of the curves and the attribution of the degree of relevance. However, from the initial base of knowledge, it is possible to generate fuzzy curves in a systematic manner. To construct those curves we organized the available data into classes and we calculated the data statistics for each class, particularly the maximum (Max), minimum (Min), mean (M) and standard deviation (Sd) 18). Using the trapezoidal form, we defined points a b c d, as shown in Figure 3, in such a way that the points in the interval a, b), by hypothesis, represent values with a maximum relevance degree equal to 1, that is:

a=M - 2*Sd, b= M + 2*Sd,

and we defined the points c d, which are the minimum and maximum values, respectively and around which the trapezoid fall is adjusted. Even though the Fuzzy System os of an imprecise nature (because it manipulates linguistic variables), it is important to point out that this very feature allows for a great ease of adjustments, especially regarding relevance functions, in such a way as to obtain the desired outputs.

***Base of Rules*:** The base of rules contains a set of rules which define the deductive nature of a Fuzzy System. The rules are applied upon the attributes, input variables, and they may be described in a simple manner by logical operators such as OR, AND (in conditional propositions and in non-conditional ones). Given a set of attributes and a classifier, we can write simples rules, such as:

if *attribute 1* AND *attribute* 2 OR *attribute3* then output1

if *attribute 1* output 2

In operations AND between two attributes the result will be the choice that has the lesser degree of relevance and in operations OR the greater degree of relevance will be chosen

***Relevance*:** Given a set of attributes for classification, inferring will consist in the parallel application of the rules through an aggregation process Azeem, 2012; Zimmermann, 2010; LIU, Feilong, 2008) which will calculate the relevance of a certain rule for the parameters of output and composition Azeem, 2012; Zimmermann, 2010; Shahjalal, 2003, which calculates the influence of each rule in the output variables – such is the deductive mechanism of a Fuzzy System.

***De-fuzzifying*:** After inference, a geometric space e is generated, called fuzzy regions. De-fuzzifying consists in mapping those regions in quantitative values expected, that is, classifying the set of input attributes. Such process may be conducted by means of several techniques: Centroid, First-of-Maximum, Middle-of-Maximum and Maximum-Criterium Azeem, 2012; Zimmermann, 2010).

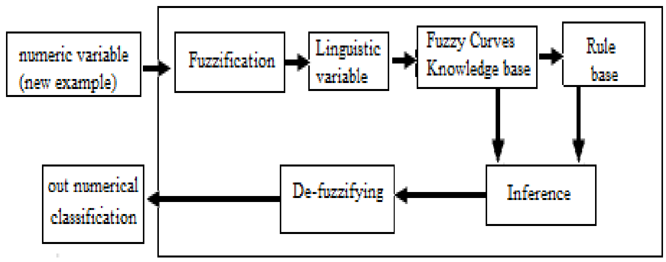


Figure 2. Schematic representation of a Fuzzy System.

It receives quantitative values, constructs the Fuzzification curves employing linguistic variables, conducts the inference of the problem, de-fuzzifies and delivers at the outcome a result in a quantitative manner. In summary, in the context of this article, a Fuzzy System receive as its input a certain set of attributes (numeric value), converts it into a linguistic domain (fuzzifying), conducts classification in the linguistic domain by means of applying a set of rules (inference) and then delivers in the output numeric values (de-fuzzifying), which, in their turn, can be mapped or adjusted for the desired classification.

Still regarding environments, as indicated in section I, this is not a work about ML or about Fuzzy Systems, but rather about the integration of two environments as it will be shown in the next section.

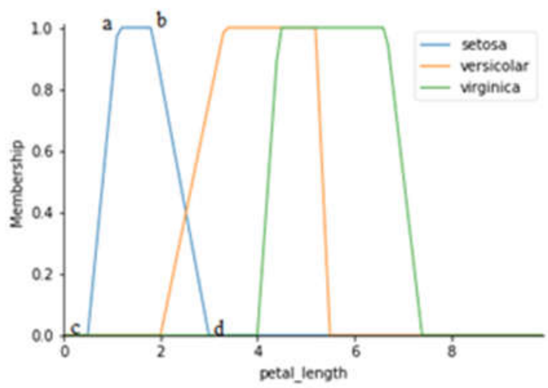


Figure 3. Examples of Fuzzy curves resulting from the fuzzifying process. Points a e b, by hypothesis, represent the interval in which the attribute has the possibility of reaching the maximum degree of relevance 1,c and d are, respectively, the points of minimum and maximum for each attribute class – for the attribute petal_length.

MATERIALS AND METHODS

Integrating The Environments

In this section, we describe a method for integrating LM and a Fuzzy System, in the context of this article. Figure 4 shows a representation of the system we propose. The sequence, in this method, is the following:

- A new example is present to LM for classification;
- ML conducts its classification;
- ML makes a call to the Fuzzy System thus providing the classified example;
- The Fuzzy System evaluates if the attributes whether the example are in a conflict region, that is, whether they are in a data region where the data are not linearly separable, which we shall call, in this article, DMZ.
- If the attributes of the example are not in the DMZ, then the Fuzzy System conducts its classification and passes it on to the ML
- If the classifications are different, then the one created by the Fuzzy System prevails and the example is inserted into the LM knowledge base.
- If the attributes of the example are in the DMZ, then the Fuzzy System does not act and the knowledge basis of the ML is not updated.

Then we observe that the Fuzzy System acts as a validator for ML under the condition that the attributes of the examples are not in the non-linearly separable region (DMZ). In section IV we describe an experiment which implements the integration we proposed in this section.

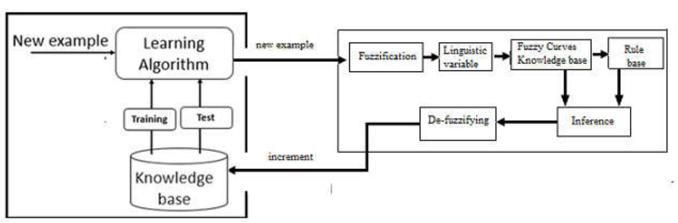


Figure 4. Schematic representation of the integration of a LM and a fuzzy system. The LM receives a new example, classifies it and then moves on to the new example for the fuzzy system make its classification and increment da knowledge base

Experiment: The experiment consists in applying the method described in section III.1 to the known iris flower classification problem

Knowledge base: The knowledge base consists of a set of 150 examples of iris flowers classified into three different species, called setosa, virginica and versicolor; each one having example samples. Classification is conducted using examples with four attributes: petal length, petal width, sepal length, sepal width. Table I shows some examples of the knowledge base

Development environment: The experiment was developed in the following environment:

Hardware: Note Book SONY VAIO Notebook, 4 GB, 2.5 GHz, 64-bit

Software: Operation System Windows 10, Framework Spyder3, Python 3.7.3, libraries: numpy, pandas 1.16.4, scipy.org sklearn-0.21.2, scikit-fuzzy 1.16.4, scikit-learn.org.

Procedure

ML Classification: An initial knowledge base is defined, with a small number of correctly classified examples. The base is small at the beginning so that we can evaluate its growth and the performance as a function of the growth. The initial base is shown in Table I and contains only 6-tuples – they are two for each class. The learning algorithm employed here is the decision tree of the sklearn library, but it could have been another algorithm available. The examples are divided into two sets: one is for training and the other one is for testing, so that it allows for measuring the performance. The other examples of the set (a total of 144) were randomly submitted (individually) to the ML, which conducted its classification.

Classifying the Fuzzy System: The Fuzzy System created for this experiment has the following basic components: a set of fuzzy curves (shown in the graph of Figure 5). The curves were generated according to the hypothesis given in the section II of this article. For the base of rules, three simple rules were defined:

```
rule1= ctrl.Rule(sepal_length['versicolor'] &
sepal_width['versicolor'] & petal_width['versicolor'] &
petal_length['versicolor'], out['versicolor'])

rule2= ctrl.Rule((sepal_length['virginica']) &
(petal_length['virginica'] & sepal_width['virginica'] &
petal_width['virginica']), out['virginica'])

rule3= ctrl.Rule((petal_length['setosa'] &
sepal_length['setosa'] & sepal_width['setosa'] &
petal_width['setosa']), out['setosa'])
```

Table 1. Set of examples of the base Iris. Four attributes and three classes

Sepal length	Sepal width	Petal length	Petal width	Class
5.1	3.5	1.4	0.2	Iris-setosa
4.9	3.0	1.4	0.2	Iris-setosa
7.0	3.2	4.7	1.4	Iris-versicolor
6.4	3.2	4.5	1.5	Iris-versicolor
6.3	3.3	6.0	2.5	Iris-virginica
5.8	2.7	5.1	1.9	Iris-virginica

Figure 5 Fuzzy Curves generated according to the section II of this article. There are four attributes, each one with a relevance function for each class (setosa, virginica and versicolor). For the Fuzzy System output, three variables were defined, each referring to one of the classes. Figure 6 shows the three possible fuzzy outputs where the classification of the example is setosa. In Figure 7, it is considered a case of DMZ, that is, the input data are not linearly separable. Even though it is possible to minimize or even eliminate this frontier region (by revising the rules or adjustments made in the degree of relevance of the attributes for each class), for the purposes of this article, we simplified the problem. In this case, by making the fuzzy system not act in updating the knowledge base of ML.

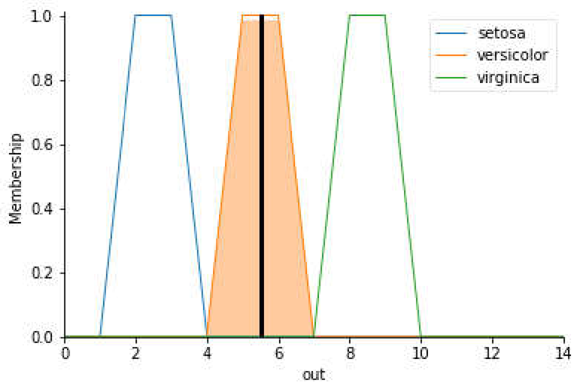


Figure 6. Fuzzy Outputs: three outputs associated to the three classes The example is classified as versicolor

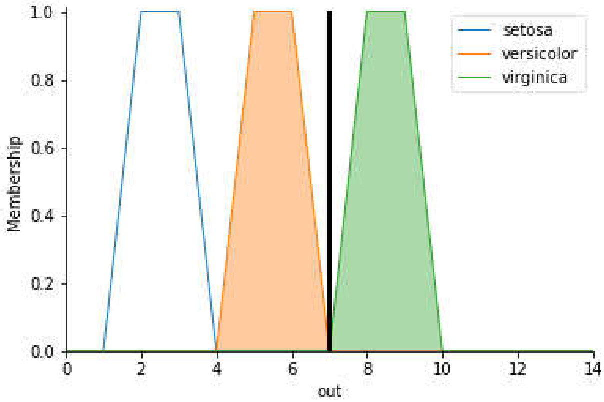


Figure 7. Frontiers conflict, also called DMZ in this article. Uncertainty to classify as versicolor or virginica. In this case, fuzzy system does not increase the base

ML integration x Fuzzy System: Having defined the environments, the next step is integration, as described in section III. We show three cases: In the first one, ML passes on the example classified by it to the Fuzzy System input, which then classifies it. If the fuzzy output is not in the DMZ and the ML classification is not equal to the fuzzy classification, the knowledge base is then incremented with the insertion of that example, a priori correctly classified by the Fuzzy System. In the first case, the Fuzzy system does not act and the ML knowledge base increment is updated with the classification conducted by the ML itself. The following are performance evaluation measures for ML and the Fuzzy System in these three cases.

RESULTS

Results are defined by two indicators, one being the ML performance as a function of the knowledge base increment and the other being the performance of the Fuzzy System when called for ML validation. The results are described below:

ML performance with base increment by means of the Fuzzy System, without DMZ. To measure ML performance as a function of the knowledge base growth, the confusion_matrix{ Narkhede, 2018; Gonçalves, 2014) metrics is employed (from sklearn library). Figure 8 shows a sketch of the confusion-_matrix, in the context of this article. It is observed the elements of the main diagonal represent the correct answers in classifying (true positive), while the other elements of the matrix show the amount of classification errors (false). Figures 9 (a) e 9(b) show the results for the size of the base equal 50 and 150, respectively. In this case, giving the ratio between correct answers (tp) and the total of evaluations – the result is 0.5 for the case 9(a) and 0.88 for case 9(b).

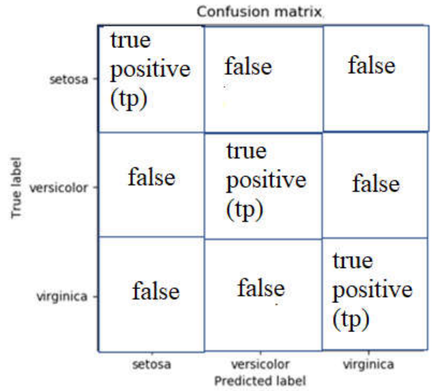


Figure 8 Structure of a confusion_matrix. The elements in the main diagonal represent the correct answers in classification (true positive) while the other elements of the matrix give the amount of classification errors (false)

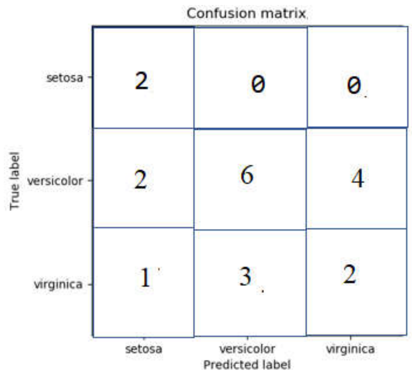


Figure 9(a). LM performance with the Fuzzy System acting, for a base size equal 50, giving the ratio between correct answers and totals, which results in 0.5.

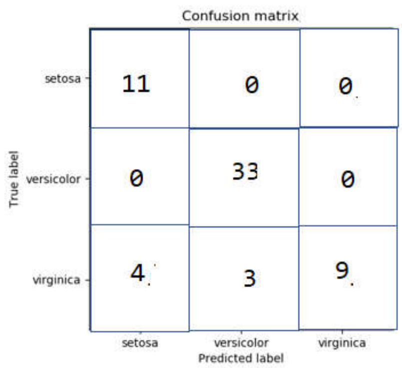
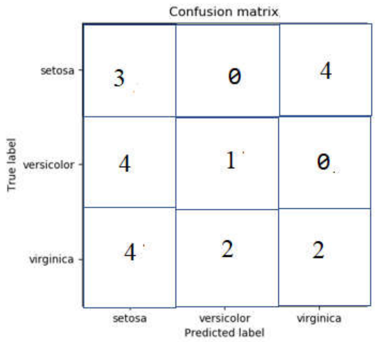


Figure 9(b). LM performance, with the Fuzzy System acting, for a base size equal 150, giving the ratio between correct answers and totals, which results in 0.88

ML performance with base increment without the Fuzzy System acting:



Figures 10 (a) and 10(b) show two performances, for a base size equal 50 and 150, respectively, without the Fuzzy System acting. In this case, the ratio between right answers and totals is given, resulting in 0.4 for the case 10(a) and 0.3 for the case 10(b).

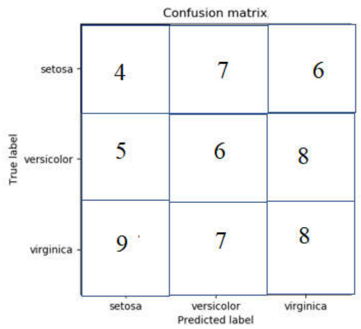


Figure 10(a). LM performance without the Fuzzy System acting, for a base size equal 50, giving the ratio between right answers and totals, which results in 0.3. Figure 10(b) LM performance without the Fuzzy System acting, for a base size equal 150, giving the ratio between right answers and totals, resulting in 0.3

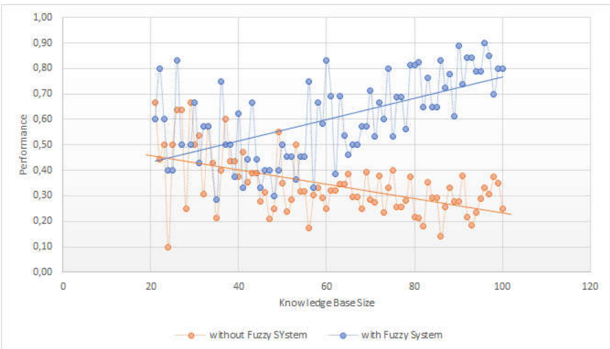


Figure 11. shows the general performance of LM, with and without the Fuzzy System acting. Figure 11 General performance of LM with an increment of the knowledge Base – with and without the Fuzzy System acting

Classifiers performance: To quantify the performance of the classifier Systems, that is, a Fuzzy System without DMZ (Iris Fuzzy) Fuzzy System with DMZ (Iris fuzzy DMZ) and wihtout Fuzzy System (Iris without Fuzzy), the measure of interest here is the number of classifications that are correct or the number of true positives (tp) versus the number of non-correct classifications, false positives (fp). For these cases, the curves of "receiver operating characteristic (ROC) Narkhede, Sarang, 2018; Gonçalves, 2014) are the ones that best suit us. ROC Curves relate the occurrences of true positives with the errors, false positives. The performance is evaluated calculating the AUC (Area Under Curve) in the normalized graph of the ROC curves. The closer to 1 the are under the curve the better the classifier's performance and the closer to 0 the area the worse its performance. Figure 12 shows the ROC curves, generated in the experiment, for the classifiers. The results show us that the classification conducted by the Fuzzy System without DMZ (Iris Fuzzy) shows an AUC close to 1, that is, 100% right answers while the one conducted by the Fuzzy System with DMZ (Iris Fuzzy DMZ) and without Fuzzy System (Iris without Fuzzy) shows an UAC around 0.5%, that is, 50% performance.

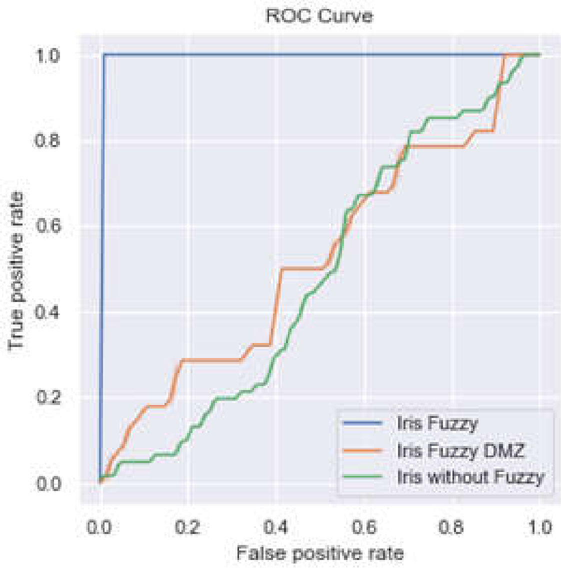


Figure 12. ROC Curves for the three cases. The Iris Fuzzy curve has an area equal to 1 (100% of right answers), while the other cases have an area around 0,5 – a performance of 50%

Analysis of results: The results show that the Fuzzy System (Iris Fuzzy) acts in the increment of the knowledge base by means of inserting correctly classified examples, as the results shown in Figure 12 and that makes the base grow and thus allow for a better performance of the ML, according to the results shown in Figure 11. The results also show that Fuzzy System does not replace ML and it does not act as a validator, but merely helps it achieving a better performance by means of the increment of its base of knowledge. The base grows with an assurance of the correctness of the examples inserted into the Fuzzy System, which then show 100% of right answers (when it acts without the conflict region, here called DMZ).

Conclusion

The results show that it is possible to automatically increment the knowledge base of a ML using the Fuzzy System in classification problems applying the approach described here. However, for each case, different adjustments are required. Even though such systems works with imprecision and subjectivity, the Fuzzy System has an advantage: it makes it easier to adjust the curves and the rules so as to obtain the best result. The conflict region, here called DMZ is emphasized to show that the Fuzzy System may be extremely simplified, even when one does not know exactly the degree of relevance of the elements in the sets of attributes. Choosing the tree as the algorithm for training was not a choice based in technical criteria (it was merely a matter of what is used more often), but actually any learning algorithm could have been used. The greater purpose o this work is to show the viability of a simpler approach to improve the performance of a ML, not to be used as a validator (which is the case of DNNs). This method was also applied to the problem of controlling environment variables (temperature, humidity, gases), with similar results.

Even though we employed a very simple base, for didactic purposes, this does not invalidate the results thus obtained by employing such approach. However, it is necessary the apply in more complex bases which, in any case, can be decomposed into simpler bases.

Acknowledgments

The results presented in this publication were obtained through Research and Development activities carried out in the Data Center Project with Internet of Things, with resources from Law no. 8.387 / 1991, and any advertising should refer to the aforementioned law according to Art. 48 of Decree 6.008 / 2006, having the company Samsung Eletrônica of the Amazônia LTDA funded this Project under the terms of said law.

REFERENCES

AGARAP, Abien Fred M. 2019. Deep Learning using Rectified Linear Units (ReLU). arXiv:1803.08375v2 cs.NE).

Azeem, M. F. 2012. Fuzzy Inference System: Theory and Applications books.google.com.

BISWAS, Saroj Kr, Bordolo Monali, Singh, H. R. 2016. Purkayastha. A Neuro-Fuzzy Rule-Based Classifier Using Important Features and Top Linguistic Features. *International Journal of Intelligent Information Technologies* (IJIIT) 12(3).

CHOLLET, François. 2018. Deep Learning with Python. Manning PUblications Co.

CPAŁKA, Krzysztof *et al*. 2013. On design of flexible neurofuzzy systems for nonlinear modelling. *International Journal of General Systems* Volume 42, Issue 6: Data Mining and Soft Computing.

GÉRON, Aurélien. Hands-on machine learning with Scikit-Learn and Tensor Flow: concepts, tools, and techniques to build intelligent systems. O'Reilly Media, Inc.", 2017

Gonçalves, Luzia. *et al*. 2014. ROC curve Estimation: an Overview. *REVSTAT: Statistical Journal*. Volume 12. Number 1. March, 1–20.

Goodfellow I. Bengio, G., Y. Courville A. 2016. Deep Learning. MIT.

HUANG *et al*. 2006. Extreme learning machine: theory and applications Neurocomputing. Elsevier.

Huang, G.B., Wang, D. H., Lan, Y. 2011. Extreme learning machines: a survey. International Journal of Machine Learning and Cybernetics, Springer, volume 2, n.2, p. 107–122.

HÜLLERMEIER, Eyke. 2015. Does machine learning need fuzzy logic? Fuzzy Sets and Systems, *Elsevier*.

Kotisiantis S.B. 2007. Supervised machine learning: A review of classification techniques. Emerging Artificial Intelligence. *Applications in Computer Engineering. IOS Press Amsterdam*, The Netherlands.2007.

LeCun, Y., Y. Bengio, G. 2015. Hinton. Deep Learning. Nature 521(7), 436-444.

LIU, Feilong. Mendel, Jerry M. 2008. Aggregation Using the Fuzzy Weighted Average as Computed by the Karnik–Mendel Algorithms. IEEE Transactions on Fuzzy Systems,

Magalhães, M.N. Lima, A. Noções de Probabilidade e Estatística. São Paulo. Edusp. 7a. ed. 2015

Metrics and scoring: quantifying the quality of predictions. https://scikitlearn.org/stable/modules/model_evaluation.htmlVisited 28/02/2020

Mitchell, T. 1997. Machine Learning. McGraw Hill, New York, NY, USA

Narkhede, Sarang. 2018. Understanding AUC – ROC Curve. Towards Data Science.

RUUSKA, Salla *et al*. 2018. Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. *Behavioural Processes* Volume 148, March Pages 56-62. Elesevier.

Shahjalal, M. 2013. Compositional rule of inference and adaptive fuzzy rule based scheme with applications. Annals of Pure and Applied Mathematics Vol. 3, No. 2, 2013, 155-168 ISSN: 2279-087X (P), 2279-0888(online) Published on 29 August.

YANN, L., Yoshua, B. 2015. Geoffrey. Deep Learning. Nature. volume 521, pages 436–444.

ZHAI, Junhai, XU., Hongyu; and LI., Yan. 2018. Fusion of Extreme Learning Machine with Fuzzy Integral. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*. VOL. 21, NO. supp02.

ZHANG, Chiyuan *et al*. 2017. Understanding deep learning requires rethinking generalization. CLR.

Zimmermann, H. J. Fuzzy set theory. WIREs Computational Statistcs. 2010

****