CORRECTION

Correction to: Co-active neuro-fuzzy inference system model as single imputation approach for non-monotone pattern of missing data

Esther-Lydia Silva-Ramirez[1] · Juan-Francisco Cabrera-Sánchez[1]

Addendum to: Neural Computing and Applications
https://doi.org/10.1007/s00521-020-05661-5

This addendum updates the information in Silva-Ramirez and Cabrera-Sánchez [5].

Many approaches have been proposed to solve missing data problem, particularly, Artificial Neural Networks (ANN) are some of the most studied approaches due to their characteristics, as being an universal approximator or their flexibility to a large amount of problems.

In this paper, Introduction section 1 shows chronologically some works that appear in the literature describing the use of different artificial neural networkbased tools to solve the problem of missing data. In a first part it is based on machine learning techniques and a second part it is based on fuzzy approach. This allows to contextualize and state the proposed research work, which presents an unexplored research.

The main originality of the paper is the novel application of CANFIS-based model to predict missing values. In particular, the CANFIS-ART model [4] is applied to predict missing values, the imputation process is performed and the obtained estimation of missing data allows to create a completed data set that improves the database quality.

The well-known CANFIS-ART model ([4] is a sophisticated neuro-fuzzy network with powerful capabilities approach which is enhanced by various algorithms which supposes the main difference to baseline CANFIS model. The algorithm Fuzzy-ART is applied to partition the input space and the training is carried out with three training

algorithms, two to the learning of parameters (Recursive Least Square and Error Backpropagation) and the third algorithm to the machine learning structure (Fuzzy-ART).

The core intent to this research is to impute both qualitative (dichotomous, ordinal, nominal) and quantitative data belonging to different fields using Coactive Neuro-Fuzzy Inference System-based method. The purpose is to impute different types of input data to obtain a completed data set which helps the difficult task of decision making. Therefore, the model is explored with databases that belong to different areas, studying its behavior. The CANFIS-based model is examined and compared with its baseline model ANN and established models (mean/mode, regression, Hotdeck), being evaluated with statistical measures such as mean of squared error or mean absolute percentage error of estimate, among others. A comparison is also made between the performance of all the models to measure the classification accuracy, which shows that an effective imputation increases the accuracy of the classifiers.

In section 2, patterns and mechanism of missing data, as well as imputation methods, are explained.

The experimental process is explained in section 3, including a brief description of the data sets, procedure perturbation, evaluation criteria and empirical experiments.

Section 4 describes the automatic procedure to impute missing values with machine learning-based techniques. General aspects of artificial neural networks and the structure of neuro fuzzy systems are introduced.

In Subsection 4.1 the neural architecture is described: a three-layer perceptron [1]. A study considering different activations functions is shown and parameters setting in Table 3 is based on previous experiments [6, 7]. Algorithm 1 not only shows the MLP algorithm, it actually shows the algorithm that describes the steps followed for the imputation process applying MLP model. The source code was implemented in Matlab R2018a, using Neural Network Toolbox [2].

In Subsection 4.2 general characteristics of the CANFIS models are commented [3], as well as CANFIS-ART

The original article can be found online at https://doi.org/10.1007/s00521-020-05661-5.

✉ Esther-Lydia Silva-Ramirez
  esther.silva@uca.es

[1] Department of Computer Science and Engineering, University of Cádiz, Avda. Universidad de Cádiz 10, 11519 Puerto Real, Cádiz, Spain

architecture of the library User's Manual [4], in particular. Aforementioned, this CANFIS-ART model uses three training algorithms, two to the learning of parameters (RLS and Error Backpropagation) and the third algorithm to the machine learning structure (Fuzzy-ART). Moreover, CANFIS-ART backpropagation algorithms [4] are described in this section.

Ad-hoc experiments were designed for the purpose of determining the values of internal parameters in CANFIS-ART model. All the parameters associated to membership function, consequence parameters, etc. are revised for each epoch during the training. To avoid overfitting in the training data an appropiate CANFIS model was searched in terms of the best generalization capability, converging to an optimal solution. The CANFIS model with the best generalization capability was obtained with the values of the parameters listed in Table 4. The results presented in this work correspond to results on the test data. Algorithm 2 not only shows the CANFIS-ART algorithm, it actually shows the algorithm that describes the steps followed for the imputation process applying CANFIS-ART model. Likewise, the source code was implemented in Matlab R2018a, using ANFIS/CANFIS Library for Simulink, developed by Konsoulas [4].

The results are discussed in section 5. In this section all imputation models are compared, measuring the models performance with statistical indexes, and measuring them in terms of their predictive accuracy using other classifiers. Moreover, this section contains a brief description of the other classification methods that are used to measure the machine learning approaches precision.

Finally, section 6 presents the main conclusions. Compared with manual process, the machine learningbased techniques considerably reduce the time and labour cost, thereby facilitating data imputation tasks. The experiments show the effectiveness that approaches based on machine learning achieve in the data cleaning process. Moreover, it is considered that the machine learning-based methods may be quite good as a supplement for other methods, being good proposals as viable alternative to the existing ones imputation models.

References

1. Bishop C (1995) Neural networks for pattern recognition. Oxford University Press, Oxford
2. Demuth H, Beale M (1997) Neural network TOOLBOX for use with matlab. User's guide. The Math Works Inc. http://www.mathworks.com
3. Jang J, Sun C, Mizutani E (1997) Neuro-fuzzy and soft computing, a computational approach to learning and machine intelligence. IEEE Trans Autom Control 42(10):1482–1484
4. Konsoulas I (2015) Adaptive neuro-fuzzy inference systems (ANFIS) library for simulink. MATLAB Central File Exchange. http://www.mathworks.com/matlabcentral/fileexchange/36098-adaptive-neuro-fuzzy-inference-systems\-anfis-library-for-simulink
5. Silva-Ramirez E, Cabrera-Sánchez J (2021) Co-active neurofuzzy inference system model as single imputation approach for non-monotone pattern of missing data. Neural Comput Appl 33:8981–9004. https://doi.org/10.1007/s00521-020-05661-5
6. Silva-Ramirez E, Pino-Mejias R, López-Coello M, Cubiles-dela-Vega M (2011) Missing value imputation on missing completely at random data using multilayer perceptrons. Neural Netw 24(1):121–129. https://doi.org/10.1016/j.neunet.2010.09.008
7. Silva-Ramirez E, López-Coello M, Pino-Mejias R (2018) An application sample of machine learning tools, such as SVM and ANN, for data editing and imputation, vol 29. Springer, Berlin, pp 259–298. https://doi.org/10.1007/978-3-319-62359-713