A Framework for the Evaluation of Machine Learning Algorithms based on Fuzzy AHP and VIKOR

Maikel-Yelandi Leyva-Vazquez[1], Francisco-Javier Poma-Torres[2], Galo-Enrique Valverde-Landivar[1] Miguel-Angel Quiroz-Martinez[1]

[1] 1Computer Science Department, Universidad Politécnica Salesiana, Guayaquil, Ecuador

ABSTRACT

Decision-makers consider various criteria, sub-criteria, alternatives, scores, and other parameters for choosing a tool. The problem is decision-making considering objectives that conflict with complex factors such as social, economic, political, technological, or environmental. It is not possible to make decisions based on a single criterion. The objective is to develop a framework for evaluating decisions based on fuzzy logic by hybridizing the fuzzy-AHP and VIKOR methods to select ML algorithms that increase these algorithms' effectiveness in specific problems. The methodology used in this proposal is oriented to exploratory research that allows to identify and define a problem or issue with a quantitative approach.

Keywords: Exemplary Paper, Human Systems Integration, Systems Engineering, Systems Modeling Language

INTRODUCTION

Machine Learning (ML) is a computational approach that includes as a funda-mental element task learning from your experience; the more experiences in-creases its measurable accuracy; the program makes decisions and achieves fore-casts set out in the data; algorithms are classified as supervised, unsupervised, semi-supervised and reinforced (Buyukozkan, G. et al. 2008, Gao, Z. et al. 2019). ML is used in engineering and medicine; the accuracy of diagnostics through ML is complex and significant to select an algorithm to make correct decisions.

This document proposes to use Analytical Hierarchy Process (AHP) and Visekriterijumsko Kompromisno Rangiranje (VIKOR) fuzzy methodology to decide between ML algorithms; below are described.

AHP is a tool for addressing subjective and objective criteria; used in the analysis of complex decisions and obtaining weights from the criteria; it is effec-tive, efficient, and straightforward; attributes are hierarchical into criteria and their subcriteria (Leyva-Vázquez, M. et al. 2014). VIKOR is a technique for optimizing complex procedures that contain compromise/conflict attributes; it focuses on selecting and sorting an option other than the options set; options are categorized through compari-sons (Buyukozkan, G. et al. 2016). There are two methods of solution; the first is synthetic evaluation such as Weighted Sum, AHP, TOPSIS, VIKOR, Probative Reasoning (ER), and Diffuse Synthetic Assessment; the second is Life Cycle Assessment (LCA).

Preliminaries

VIKOR method (Gao, Z. et al. 2019, San Cristóbal, J.R. et al. 2011). is based on the classification and selection of the set of alternatives through multicriteria; the rating index that measures proximity to the best resource (alternative) is used; the function of each criterion evaluates each alternative, after classification the best resource is checked

The VIKOR classification algorithm defines the following stages:

Step 1: It must be determined the high values of f_i^* and the low values of f_i^- in the criteria, i.e. Dataset, Data Preprocessing, Area Impact, Training and Evaluation *(i = 1,2,…n)*; if the i-th condition is an optimal benefit then $f_i^* = max_j f_{ij}$ and $f_i^- = min_j f_{ij}$; if the i-th condition is a low benefit, then $f_i^* = min_j f_{ij}$ and $f_i^- = max_j f_{ij}$;

Step 2: It must be calculated the values of S_j and R_j through relationships:

$$S_j = \sum_{i=1}^{n} W_i\left(f_i^* - f_{ij}\right)/(f_i^* - f_i^-) \quad (1)$$

$$R_j = max_i\left[W_i\left(f_i^* - f_{ij}\right)/(f_i^* - f_i^-)\right] \quad (2)$$

Here: S_j is the measure of usefulness, R_j is the measure of repentance, W_i is the weight of the condition i-th; the resource generated by min_jS_j is a "majority utility"; the resource generated by min_jR_j is a particular minimum of repentance from the opponent.

Step 3: Calculate index Q_j:

$$Q_j = \frac{v\left(S_j - S^*\right)}{(S^- - S^*)} + (1 - v)\frac{\left(R_j - R^*\right)}{(R^- - R^*)} \quad (3)$$

Here: $S^*=min_jS_j$, $S^-=max_jS_j$, $R^*=min_jR_j$, $R^-=max_jR_j$, S_j and R_j were calculated in the last step; v is the strategy as the "majority utility", value of v=0.5.

Step 4: Here value of Q is sorted in descending order; the ML algorithm of the priority of alternatives is proposed as a solution (Ray, S. et al. 2019), it is the best classification by the measure Q (Minimum).

Application of AHP-VIKOR Hybrid MCDM Approach.

Decision analysis is a discipline that aims to help decision-makers reach a reliable decision; combining methods helps increase the reliability of this process. AHP is a method that can be used in the weighting of criteria according to the values defined by Saaty. This AHP procedure scheme uses the Saaty ratio scale [8] (Ren, Z. et al. 2019) from 1 to 9 represented in Table 1.

Table 1. Ratio scale 1-9.

Scale	1/9	1/8	1/7	1/6	1/5	1/4	1/3	1/2	1	2	3	4	5	6	7	8	9
Priority grade	Smaller scale, lower priority								Equal	Larger scale, higher priority							

The Fuzzy-AHP process proves to be a useful technique for multicriteria deci-sion-makers in fuzzy environments and has had a large number of applications in recent years.

Methods

The methodology used in this proposal is oriented toward exploratory research that allows to identification and defines a problem or issue with a quantitative approach. A common decision-making scheme consists of the following phases: Framework, Information gathering, Evaluation of the alternatives, and Selection of alternatives.

During the frame-of-reference phase, the structures and elements of the decision problem are defined. Experts provide information according to the defined framework. The AHP method is used for the weighting of the criteria using fuzzy logic to represent uncertainty. From these weights, VIKOR has used selection of alternatives to carry out a resolution process to calculate the evaluations for the set of alternatives, using aggregation operators.

1 Case study

Next, a case study is shown to demonstrate the applicability of the proposal. The following criteria were defined:
Data Set evaluates the source of the data; the subcriteria data types grade the data found in text or image; the source subcriteria data qualifies that deliver, like governments, universities, foundations, hospitals or third parties; the subcriteria recovery site checks if they are reliable places like databases or files in free format.

Data preprocessing contains the activities to be performed on the dataset once obtained; sometimes, it is necessary to clean up the data before the ML algorithms process it.

Impact on the area determines the weight in utilization in an area that you want to apply the ML algorithm; the subcriteria algorithms used in the area; the subcriteria amount of workouts describes the times the algorithm used the data before use in testing; the subcriteria amount of the available dataset describes the availability of data for algorithm training; subcriteria available programs express how many tools exist to run on the dataset; subcriteria available platforms expresses the operating systems that support the available programs.

Training evaluates the requirements that the algorithm needs to be executed on a platform; subcriteria amount of data are the records or rows of the file or number of images; subcriteria amount of features in the data are the attributes that the set possesses; subcriteria training times are the hours, minutes and seconds that the algorithm takes to run; subcriteria number of parameters is related to the activation functions of the algorithm.

Evaluation Criteria contains the quantitative indicators that apply to the ML algorithm; Subcriteria Accuracy is the quality score of the ML algorithm in classification tasks; Recall subcriteria determines the amount that the ML algorithm can identify; subcriteria F1 score is a combination of precision measures and recall in a single value, provides a comparison of performance.

Among the five proposed alternatives, it is based on the score of tests performed on ML algorithms; the alternatives are ANN, SVM, LR, Random Forest y K Nearest Neighbor (Fig. 1).

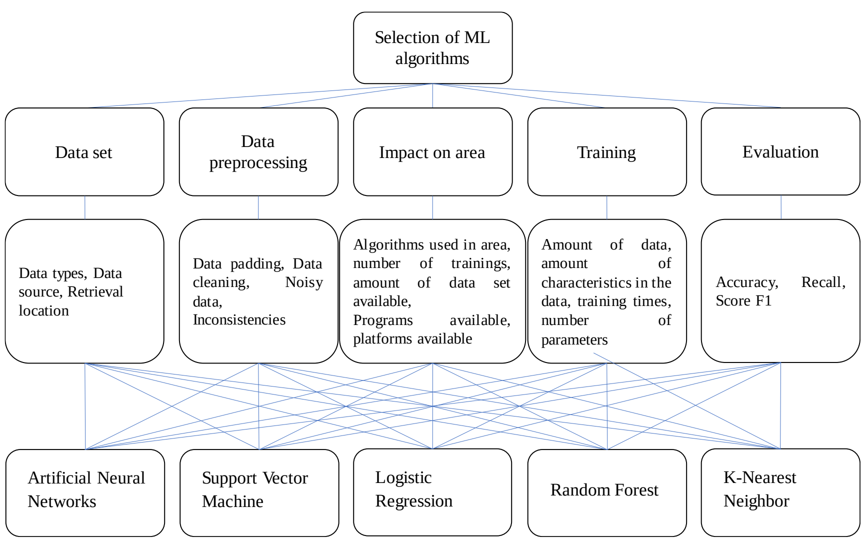

Fig. 1. Hierarchy for multicriteria decision making.

The preference for weighting the criteria assigned by the relative weights of the attributes is through the Fuzzy AHP scale found in Table 1; Table 2 shows the relative importance used in this project.

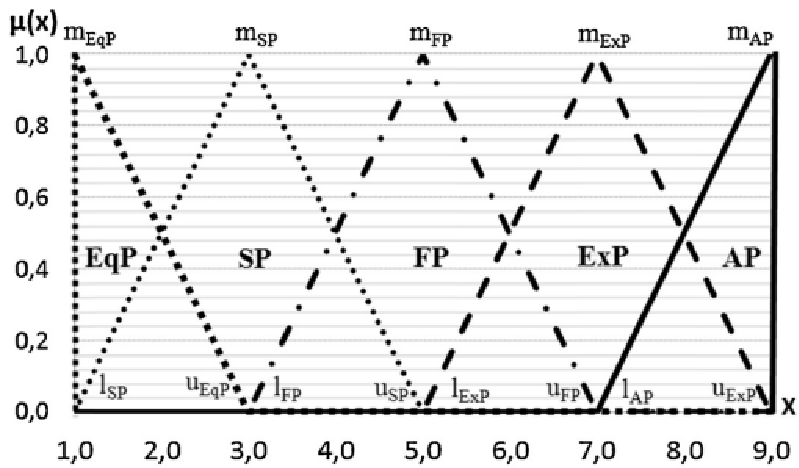

Fig. 2. The comparative linguistic scale of the weights of the judgments and appreciations of the alternatives.

Conclusions

It was concluded that the proposed framework is a technique for determining the detailed evaluation of ML algorithms and resolving multicriteria decision-making; this approach is a tool that uses AHP in assigning weight to criteria, and these scores are processed with VIKOR in multicriteria optimization; and get an alternative/compromise solution, here the evaluation is scientific, fair, and rational.

ML algorithms measure qualitative characteristic data and the results obtained from the data are quantitative; this proposal aims to get the best algorithm for measuring data. In future computing with words with the model based on 2- tuples, to increase the interpretability of the results will be used.

Acknowledgments. This work has been supported by the GIIAR research group and the Universidad Politécnica Salesiana.

REFERENCES

Buyukozkan, G., Feyzioglu, O., Gocer, F.: Evaluation of hospital web services using intuitionistic fuzzy AHP and intuitionistic fuzzy VIKOR. In: 2016 IEEE (IEEM). pp. 607–611. IEEE (2016)

Gao, Z., Liang, R.Y., Xuan, T.: VIKOR method for ranking concrete bridge repair projects with target-based criteria. Results Eng. 3, 100018 (2019). https://doi.org/10.1016/j.rineng.2019.100018

Leyva-Vázquez, M., Quiroz-Martínez, M.A., Portilla-Castell, Y., Hechavarría-Hernández, J.R., González-Caballero, E.: A New Model for the Selection of Information Technology Project in a Neutrosophic Environment. Neutrosophic Sets Syst. 344

Leyva-Vázquez, M., Pérez-Teruel, K., John, R.I.: A model for enterprise architecture scenario analysis based on fuzzy cognitive maps and OWA operators. In: CONIELECOMP 2014 - 24th International Conference on Electronics, Communications and Computers (2014)

Lihong, M., Yanping, Z., Zhiwei, Z.: Improved VIKOR algorithm based on AHP and Shannon entropy in the selection of thermal power enterprise'|s coal suppliers. Proc. Int. Conf. Inf. Manag. Int. Conf. Inf. Manag. Innov. Manag. Ind. Eng. ICIII 2008. 2, 129–133 (2018). https://doi.org/10.1109/ICIII.2008.29

Ray, S.: A Quick Review of Machine Learning Algorithms. In: 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). pp. 35–39. IEEE (2019)

Ren, Z., Xu, Z., Wang, H.: The Strategy Selection Problem on Artificial Intelligence With an Integrated VIKOR and AHP Method Under Probabilistic Dual Hesitant Fuzzy Information. IEEE Access. 7, 103979–103999 (2019). https://doi.org/10.1109/ACCESS.2019.2931405

San Cristóbal, J.R.: Multi-criteria decision-making in the selection of a renewable energy project in spain: The Vikor method. Renew. Energy. 36, 498–502 (2011). https://doi.org/10.1016/j.renene.2010.07.031

Tian, G., Zhou, M., Zhang, H., Jia, H.: An integrated AHP and VIKOR approach to evaluating green design alternatives. In: 2016 IEEE 13th (ICNSC). pp. 1–6. IEEE (2016)

Yu, X., Zhang, H., Bouras, A., Ouzrout, Y., Sekhari, A.: Multi-Criteria Decision Making for PLM Maturity Analysis based on an Integrated Fuzzy AHP and VIKOR Methodology. J. Adv. Manuf. Syst. 17, 155–179 (2018). https://doi.org/10.1142/S0219686718500105