Research Statement

Fatemeh Aghaeipoor

1. Overview

My research interests lie primarily in developing Machine Learning (ML) models in the light of XAI, eXplainable Artificial Intelligence[9], the new resurgence of Artificial Intelligence that attempts to provide more explainable, interpretable, and transparent models. Although there is a little consensus on these terms in the literature [7], all have a common sense to clarify the inner functionality of the systems for direct interaction with human cognition. Indeed, these characteristics enable users to better understand, trust, modify, and manage the behaviors of intelligent systems and ML models. The success of transparent systems is specifically highlighted in those applications in which the decision-making or decision-support processes should be made clear to the users.

The objective of my research is to develop a set of quantitative and qualitative machine learning approaches and building well-interact and easy-understanding data-based models. To this end, I have been following two interrelated tracks: boosting explainability in the inherently transparent models or alleviating the crucial weaknesses in the black-box models. In the following, I describe my scientific works and my plans to further pursue these exciting directions.

2. Current Research: XAI and Fuzzy Modeling

There are many fields of application where it is necessary, if not essential, to give an explanation of the phenomenon under study. It is no longer enough to simply apply a ML method to extracts the models, but they must be comprehensible in order to provide real decision support systems. For this reason, a strong movement has emerged in favour of the eXplainable Artificial Intelligence that aims to respond to the "how" and "why" of the operation of automatic models.

In this context, I began my research career by developing interpretable fuzzy models based on the available (Big) data of different applications. Indeed, fuzzy models owing to their inherent characteristics in the use of linguistic fuzzy labels, which contain the semantic knowledge inspired by the human language, are straightforward transitions to globally explain the phenomena to the practitioner. In these systems, we can effectively manage the involved factors, leading to efficient and straightforward predictive models. These systems have applications in a wide variety of disciplines, such as bio-medical, fraud detection, power consumption, and financial problems. The majority of my previous researches took advantage of evolutionary fuzzy learning. Some of my previous works in the interpretability of sequential models are as follows:

- In [4], we developed an interpretable multi-objective evolutionary fuzzy system for learning high-dimensional regression data, in which the focus was on automatic learning of the available data with the priority of being easily handled for human mind processing.
- [1], [3], and [5] were also developed to improve the interpretability and accuray of these systems from diffrent points of veiw.

Recently, in collaboration with some of my colleagues at the University of Granada, we attempted to boost the explainability of fuzzy systems in the context of Big Data and distributed algorithms [2, 6]. However, in the case of addressing Big Data analytics, providing explainable yet scalable models is still an open issue, which I am highly motivated to work on.

3. Future Research: ML and Big Data under the umbrella of XAI

ML models are increasingly applied to solve complex and computational problems of human life. They are establishing intelligent systems perceiving, learning, deciding, and operating almost without human intervention. In such a situation where these intelligent systems are highly employed in critical aspects of our lives like medicine, law, finances,

self-driving cars, robotic assistants, and so on, understanding and explaining their internal logic finds a significant importance [9], as it helps human users to trust sincerely, manage effectively, avoid biases, evaluate decisions, and provide more robust ML models [8]. All these objectives are nowadays following in the light of XAI [7].

Looking forward, I will continue my research to bridge between fundamental XAI and applied ML. It needs to work on the human factors of this technology and can be investigated considering two directions: better understanding of the human interactions and building transparent systems providing rich interactions. Towards this goal, I anticipate the following future directions: developing novel machine learning models with the priorities of XAI, boosting the available methods regarding their inner transparency, design efficient user interfaces, integration the capabilities of natural language processing and cognitive science into Ml tools, targeting XAI in many application domains like biomedical, text analysis, citizen science, the credibility of information, and so on. Moreover, there are still several open theoretical issues in this context, like defining effective local and/or general explainability measures in order to quantify and qualify models of the same/different paradigm(s).

Focusing on Big Data analytics, creating explainable ML models may be involved with some intrinsic challenges. They are mostly computationally intensive. Even if they can come up with scalable solutions, they may be complex in terms of interpreting their distributed manner. I am extremely passionate about working on this area of research.

Finally, given my background in ML and my programming skills, I very much like to do research in different aspects of XAI and ML with both applied and theoretical contributions. In this regard, I am willing to collaborate with other researchers mainly in applied Ml and Big Data, which gives me a great opportunity to expand my expertise as well as addressing the important problems of our society intelligently.

References

[1] F. Aghaeipoor and M. Eftekhari. EEFR-R: extracting effective fuzzy rules for regression problems, through the cooperation of association rule mining concepts and evolutionary algorithms. *Soft Computing*, 23(22):11737–11757, 2019.

[2] F. Aghaeipoor, M. M. Javid, I. Triguero, and A. Fernández. Chi-BD-DRF: Design of scalable fuzzy classifiers for big data via a dynamic rule filtering approach. In *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–7, 2020.

[3] F. Aghaeipoor and M. Javidi. A hybrid fuzzy feature selection algorithm for high-dimensional regression problems: an mRMR-based framework. *Expert Systems with Applications*, 162:113859, 2020.

[4] F. Aghaeipoor and M. M. Javidi. MOKBL+ MOMs: An interpretable multi-objective evolutionary fuzzy system for learning high-dimensional regression data. *Information Sciences*, 496:1–24, 2019.

[5] F. Aghaeipoor and M. M. Javidi. On the influence of using fuzzy extensions in linguistic fuzzy rule-based regression systems. *Applied Soft Computing*, 79:283–299, 2019.

[6] F. Aghaeipoor, M. M. Javidi, and A. Fernández. IFC-BD: An interpretable fuzzy classifier for boosting explainable artificial intelligence in big data. *IEEE Transactions on Fuzzy Systems*, In Press, 2021.

[7] A. B. Arrieta, N. Díaz-Rodríguez, J. D. Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58:82 – 115, 2020.

[8] R. Marcinkevics and J. Vogt. Interpretability and explainability: A machine learning zoo mini-tour. *ArXiv*, arXiv:2012.01805, 2020.

[9] T. Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267:1–38, 2019.