Introduction to Fairness in Algorithmic Decision Making mini-track

Boris Delibašić
University of Belgrade
boris.delibasic@fon.bg.ac.rs

Sandro Radovanović
University of Belgrade
sandro.radovanovic@fon.bg.ac.rs

Abstract

A vast application of machine learning and decision-making algorithms for decision support in various areas of life caused the need for the algorithms to take into account additional constraints, such as non-discriminatory behavior or imposing fairness, or social welfare prior to proposing decisions to decision makers. These constraints can be fulfilled by carefully guiding the whole decision-making and data governance process, by adjusting decision-making, data mining and machine learning algorithms to fulfill additional constraints. For example, by adapting CRISP-DM methodology to account for possible biases, by imposing instance-dependent cost-sensitive learning, or enforcing equality in data envelopment analysis as presented in this mini-track.

1. Introduction

The Fairness in algorithmic decision making mini-track addresses topics related to imposing fairness requirements and conditions in algorithmic decision making. With the introduction of regulations such as General Data Protection Regulation, European Commission Artificial Intelligence Act, and Algorithmic Accountability Act, algorithms used for automatization of human decision-making in areas such as classification, recommendation, ranking, are subject to non discriminatory behavior. [1]

Although algorithms are not a recent invention, they are being increasingly used in many systems to support business decision making. Such systems often rely on a large amount of data to extract knowledge for more informed decisions. Even with human intervention, the impact of the decision on people can be significant, such as access to education, employment, medical treatment etc. [2]. However, allowing algorithmic decision making tools to make or influence decisions raises ethical, legal, and technical issues. If these issues are neglected, benefits of such systems will result in discrimination, unfair practices, manipulations etc. To address and mitigate these issues one needs to design and create a mechanism or technique to achieve algorithmic fairness. Fairness can be usually imposed in algorithmic decision making prior to algorithm learning or application (pre-processing) [3, 4], during the decision making model learning (in-processing) [5, 6, 7], or prior to model application (post-processing) [8]. Besides this, one must observe short-term and long-term effects of algorithm interventions and design fair decision making mechanisms [9].

One such mechanism is to impose new standards through the whole data mining process. Cross-industry standard process for data mining (CRISP-DM) is an industry and technology independent model for organizing ML projects' development. The model still lacks fairness concerns related to ML technologies. To address this important theoretical and practical gap in the literature a new model, Fair CRISP-DM, which categorizes and presents the relevant fairness challenges in each phase of project development is proposed [10].

Decision-making algorithms are, besides machine learning algorithms, at the forefront of fair algorithmic decision making. Data envelopment analysis (DEA) is one of the classical decision-making algorithms concerned with calculating efficiency scores of decision-making units. Here, a MAX-MIN fair cross-efficiency data envelopment analysis (DEA) model that solves the problem of high variance cross-efficiency scores is proposed. The MAX-MIN cross-efficiency procedure is in accordance with John Rawls's Theory of justice by allowing efficiency and cross-efficiency estimation such that the greatest benefit of the least-advantaged decision making unit is achieved. The proposed mathematical model is tested on a healthcare related dataset. The results suggest that the proposed method solves several issues of cross-efficiency scores. First, it enables full rankings by having the ability to discriminate between the efficiency scores of DMUs. Second, the variance of cross-efficiency scores is reduced, and finally, fairness

URI: https://hdl.handle.net/10125/79520
978-0-9981331-5-7

is introduced through optimization of the minimal efficiency scores. [11]

Finally, schemes of learning classification algorithms are being tested. Traditionally, classification algorithms aim to minimize the number of errors. This approach can lead to sub-optimal results for the common case where the actual goal is to minimize the total cost of errors and not their number. To address this issue, a variety of cost-sensitive machine learning techniques have been suggested. Methods have been developed for dealing with both class- and instance-dependent costs. Using instance-dependent costs instead of class-dependent costs leads to improved performance for cost-sensitive performance measures, but worsens performance for cost-insensitive metrics. These results confirm that instance-dependent methods are useful for many applications where the goal is to minimize costs. [12]

References

[1] D. Abu-Elyounes, "Contextual fairness: A legal and policy analysis of algorithmic fairness," *University of Illinois Journal of Law, Technology Policy*, pp. 1 – 54, 2020.

[2] S. Fazelpour and D. Danks, "Algorithmic bias: Senses, sources, solutions," *Philosophy Compass*, vol. 16, no. 8, p. e12760, 2021.

[3] M. Feldman, S. A. Friedler, J. Moeller, C. Scheidegger, and S. Venkatasubramanian, "Certifying and removing disparate impact," in *proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 259–268, 2015.

[4] S. Rančić, S. Radovanović, and B. Delibašić, "Investigating oversampling techniques for fair machine learning models," in *International Conference on Decision Support System Technology*, pp. 110–123, Springer, 2021.

[5] A. Petrović, M. Nikolić, S. Radovanović, B. Delibašić, and M. Jovanović, "Fair: Fair adversarial instance re-weighting," *arXiv preprint arXiv:2011.07495*, 2020.

[6] S. Radovanović, A. Petrović, B. Delibašić, and M. Suknović, "Enforcing fairness in logistic regression algorithm," in *2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, pp. 1–7, IEEE, 2020.

[7] A. Petrović, M. Nikolić, M. Jovanović, M. Bijanić, and B. Delibašić, "Fair classification via monte carlo policy gradient method," *Engineering Applications of Artificial Intelligence*, vol. 104, p. 104398, 2021.

[8] G. Pleiss, M. Raghavan, F. Wu, J. Kleinberg, and K. Q. Weinberger, "On fairness and calibration," *arXiv preprint arXiv:1709.02012*, 2017.

[9] S. Mitchell, E. Potash, S. Barocas, A. D'Amour, and K. Lum, "Algorithmic fairness: Choices, assumptions, and definitions," *Annual Review of Statistics and Its Application*, vol. 8, pp. 141–163, 2021.

[10] V. Singh, A. Singh, and K. Joshi, "Fair crisp-dm: Embedding fairness in machine learning (ml) development life cycle," *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022.

[11] S. Radovanović, B. Delibašić, A. Marković, and M. Suknović, "Achieving max-min fair cross-efficiency scores in data envelopment analysis," *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022.

[12] T. Vanderschueren, W. Verbeke, B. Baesens, and T. Verdonck, "nstance-dependent cost-sensitive learning: do we really need it?," *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022.