.

Leveraging Time-Aware Causal Algorithmic Recourse for Promoting Fairness*

Isacco Beretta[1], Martina Cinquini[1,*]

[1]*Department of Computer Science, University of Pisa, Pisa*

Abstract

Algorithmic Recourse (AR) addresses adverse outcomes in automated decision-making by offering actionable recommendations. However, current state-of-the-art methods overlook the interdependence of features and do not consider the temporal dimension. To fill this gap, TIME-CAR emerges as a pioneering approach that integrates temporal information. Building upon this formulation, this work investigates the context of fairness, specifically focusing on the implications for marginalized demographic groups. Since long wait times can significantly impact communities' financial, educational, and personal lives, exploring how time-related factors affect the fair treatment of these groups is crucial to suggest potential solutions to reduce the negative effects on minority populations. Our findings set the stage for more equitable AR techniques sensitive to individual needs, ultimately fostering fairer suggestions.

Keywords

Algorithmic Recourse, Fairness, Consequential Recommendations

1. Introduction

Algorithmic Recourse (AR) aims to provide actionable recommendations to reverse negative outcomes from automated decision-making systems [1]. Recent advancements in AR have incorporated causality to ensure plausible counterfactuals and align with human cause-effect reasoning [2, 3, 4]. However, a limitation of these methods is their inability to integrate the temporal dimension, which can lead to erroneous identification of effective features in terms of cost and time. In [5] is presented TIME-CAR, the first proposal on integrating the temporal dimension into a Causal AR problem. In this paper, we investigate the implications of fairness within the TIME-CAR framework, focusing on how longer periods needed for certain tasks affect marginalized demographic groups and their connection to socioeconomic factors. The work aims to formulate fairer AR methods that are sensitive to these populations' unique needs and time constraints.

Background Formally, the AR problem can be formulated in the following terms: given a binary classifier $h : \mathbf{X} \rightarrow \{0, 1\}$, and an instance X for which $h(X) = 0$, the goal is to select the action $\mathbb{A}_{\delta^*}$ satisfying

$$\mathbb{A}_{\delta^*} = \arg\min_{\mathbb{A}_\delta} c(X, \mathbb{A}_\delta) \quad s.t. \quad h(\mathbb{A}_\delta(X)) = 1 \tag{1}$$

where $\mathbb{A}_\delta(X)$ is a modified version of X.

In [5] is presented a cost function that incorporates the temporal dimension: $c(X, Y, \mathbb{A}_\delta) = c_s(X, \mathbb{A}_\delta) + \lambda c_t(X, \mathbb{A}_\delta, Y)$, where X is the individual's initial state, Y is the target state, and $\mathbb{A}_\delta$ is the action taken to obtain the transition between them. c_s is the feature space cost, c_t is the time cost, and λ is a tunable parameter that values how important is time compared to the other features. Time-unaware algorithms set $\lambda = 0$.

Discovery Science - Late Breaking Contributions 2024

*Corresponding author.
isacco.beretta@phd.unipi.it (I. Beretta); martina.cinquini@phd.unipi.it (M. Cinquini)
https://marti5ini.github.io/ (M. Cinquini)
0000-0002-0463-6810 (I. Beretta); 0000-0002-0877-7063 (M. Cinquini)

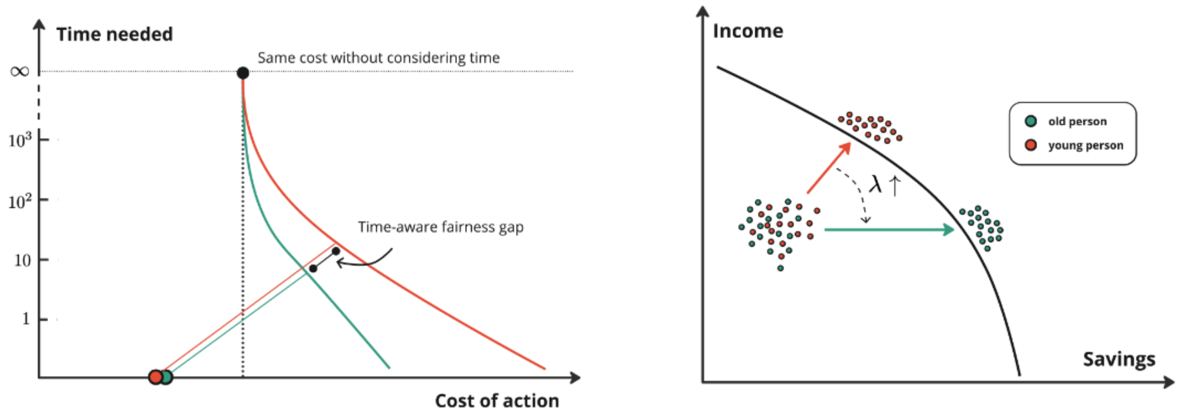

Figure 1: Possible scenarios where the time cost matters from a fairness perspective. (Left). Not accounting for time could introduce hidden biases in recommendation algorithms. (Right). Let $cost(\delta) = c_s(\delta) + \lambda \mathbb{1}(\delta_{income} \neq 0)$, it could be that λ is depending on age.

2. Fairness through Time-Aware Recourse

Same cost, different times. The temporal aspect frequently plays a crucial role in assessing the appropriateness and effectiveness of advice given to individuals by automated decision-making systems. We explore scenarios where $A \subset X$ represents sensitive attributes among the features, and we examine the case of two individuals, i_1 and i_2, with different sensitive attributes' values, such that $A(i_1) \neq A(i_2)$. We hypothesize that the cost recommendations from the time-unaware automatic decision system for these individuals, $c_s(X(i_1), \mathbb{A}_{\delta_1})$ and $c_s(X(i_2), \mathbb{A}_{\delta_2})$, are approximately equal. Despite the difference in sensitive attributes, the system suggests similar cost interventions for both individuals. However, the temporal cost c_t could vary significantly between them, meaning one individual might need more time to achieve the desired state than the other, see Fig. 1 (Left).

Not everyone values time equally. Time may be regarded as a resource whose value varies based on individual factors. Fig. 1 (Right) shows this idea through applying for a loan scenario. The value of λ might be higher for the older population as they are likely closer to retirement and have a limited window to recover from financial setbacks. Conversely, younger individuals might have a lower λ value given their longer time horizon to adjust their savings behavior. Financial models must be calibrated to accommodate varying λ values across different demographic segments. This knowledge enables the design of customized recommendations sensitive to each individual's dynamics and the time-related evaluation of changes within their specific societal and economic contexts.

Actionability as a time constraint. Actionability is considered one of the crucial aspects in a counterfactual generation process [6]. We propose expanding the concept beyond the notion of *being able to act upon* to include *the ability to do so within a reasonable timeframe.* Indeed, if the action required to implement a recommendation is excessively time-consuming or impractical, the recommendation becomes unhelpful for the user. In the constrained optimization framework of AR, the actionability threshold is directly controlled by the maximum time constraint, denoted as max_{time}. This parameter can be determined a priori or *adapted based on the user's specific requirements each time a request is made.* In the latter case, the parameter enables personalized control over actionability for the applicant. From this perspective, we propose a new interpretation of *fair* recommendations expressed as follows: *a time-aware algorithmic recourse model is fair if its recommendations remain fair under any fixed time constraint.*

Acknowledgments

This work has been partially supported by the EU H2020 programme under the funding schemes ERC-2018-ADG G.A. 834756 "XAI: Science and technology for the eXplanation of AI decision making", "SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics", by the European Commission under the NextGeneration EU programme – National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR) – Project: "SoBigData.it – Strengthening the Italian RI for Social Mining and Big Data Analytics" – Prot. IR0000013 – Avviso n. 3264 del 28/12/2021, and M4C2 - Investimento 1.3, Partenariato Esteso PE00000013 - "FAIR - Future Artificial Intelligence Research" - Spoke 1 "Human-centered AI", and by Fondo Italiano per la Scienza FIS00001966 MIMOSA.

References

[1] S. Wachter, B. Mittelstadt, C. Russell, Counterfactual explanations without opening the black box: Automated decisions and the gdpr, Harv. JL & Tech. 31 (2017) 841.

[2] A. Karimi, B. Schölkopf, I. Valera, Algorithmic recourse: from counterfactual explanations to interventions, in: M. C. Elish, W. Isaac, R. S. Zemel (Eds.), FAccT '21: 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event / Toronto, Canada, March 3-10, 2021, ACM, 2021, pp. 353–362.

[3] A.-H. Karimi, J. Von Kügelgen, B. Schölkopf, I. Valera, Algorithmic recourse under imperfect causal knowledge: a probabilistic approach, Advances in neural information processing systems 33 (2020) 265–277.

[4] J. Pearl, D. Mackenzie, The book of why: the new science of cause and effect, Basic books, 2018.

[5] I. Beretta, M. Cinquini, The importance of time in causal algorithmic recourse, in: World Conference on Explainable Artificial Intelligence, Springer, 2023, pp. 283–298.

[6] R. Guidotti, Counterfactual explanations and how to find them: literature review and benchmarking, Data Mining and Knowledge Discovery (2022) 1–55.