SafeAR: Towards Safer Algorithmic Recourse by Risk-Aware Policies

Haochen Wu
haochenw@umich.edu
University of Michigan, Ann Arbor
Michigan, USA

Shubham Sharma
shubham.x2.sharma@jpmchase.com
J.P. Morgan AI Research
New York, USA

Sunandita Patra
sunandita.patra@jpmchase.com
J.P. Morgan AI Research
New York, USA

Sriram Gopalakrishnan
sriram.gopalakrishnan@jpmchase.com
J.P. Morgan AI Research
New York, USA

ABSTRACT

With the growing use of machine learning (ML) models in critical domains such as finance and healthcare, the need to offer recourse for those adversely affected by the decisions of ML models has become more important; individuals ought to be provided with recommendations on actions to take for improving their situation and thus receive a favorable decision. Prior work on sequential algorithmic recourse—which recommends a series of changes—focuses on action feasibility and uses the proximity of feature changes to determine action costs. However, the uncertainties of feature changes and the risk of higher than average costs in recourse have not been considered. It is undesirable if a recourse could (with some probability) result in a worse situation from which recovery requires an extremely high cost. It is essential to incorporate risks when computing and evaluating recourse. We call the recourse computed with such risk considerations as Safer Algorithmic Recourse (SafeAR). The objective is to empower people to choose a recourse based on their risk tolerance. In this work, we discuss and show how existing recourse desiderata can fail to capture the risk of higher costs. We present a method to compute recourse policies that consider variability in cost and connect algorithmic recourse literature with risk-sensitive reinforcement learning. We also adopt measures "Value at Risk" and "Conditional Value at Risk" from the financial literature to summarize risk concisely. We apply our method to two real-world finance related datasets and compare policies with different levels of risk-aversion using risk measures and recourse desiderata (sparsity and proximity).

CCS CONCEPTS

• **Computing methodologies** → **Machine learning**; *Reinforcement learning*; • **Human-centered computing**;

XAI-FIN-2023, November 27, 2023, New York, USA

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

KEYWORDS

Algorithmic Recourse, Explainable AI, Reinforcement Learning

ACM Reference Format:
Haochen Wu, Shubham Sharma, Sunandita Patra, and Sriram Gopalakrishnan. 2018. SafeAR: Towards Safer Algorithmic Recourse by Risk-Aware Policies. In *Proceedings of Workshop on Explainable AI in Finance (XAI-FIN-2023).* ACM, New York, NY, USA, 15 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Machine learning (ML) models are increasingly being used to make decisions in a wide array of financial applications such as loan approvals [25] and Portfolio Optimization [1]. Given their impact on society, the importance of algorithmic recourse has increased [42]. Algorithmic recourse refers to a computed recommendation provided to an end user which suggests specific changes they can make to convert an unfavorable outcome (e.g., loan rejection), into a favorable one. For a recourse to be helpful, the suggested change ought to be actionable; for example, one can change their savings balance but not their age. Existing recourse work has considered the cost of taking the recommended actions [31, 43, 44]. However, they do not consider the risk of higher costs. In this work, *risk* means the potential for higher costs during the recourse due to the possibility of reaching adverse states; this can happen due to uncertainties in action effects (not deterministic). The cost could be in terms of time required, effort, financial resources, etc. Without incorporating risks—which is ignoring uncertainties or only minimizing the expected costs—the recipient of an algorithmic recourse may be caught unaware and unprepared for situations with high costs. By offering recourse policies with risk measures, we can help people be aware of how much risk is involved in each policy and choose a safer one. The *recourse policy* in this work refers to the recommended actions for all possible states a person might encounter, as opposed to a single deterministic sequence of actions.

To further understand the need for risk considerations, let's look at the existing algorithmic recourse approaches that use *counterfactual explanation* (CE) methods to give recourse recommendations. CE methods find "*the most similar instances to the feature vector describing the individual* (their initial state), *that also get the desired prediction from the model*" [24]. The assumption is that minimizing feature-space differences translates to a recourse that requires less cost to reach the desired outcome. Rather than providing a single

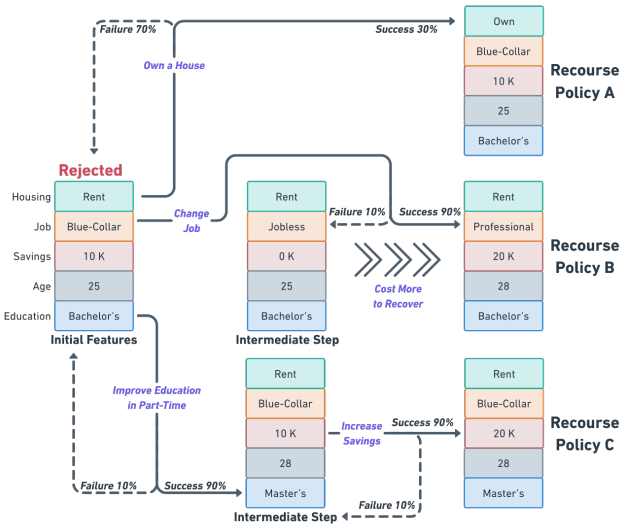

Figure 1: Recourse policies for credit loan approvals. Policy A has only one feature change (low sparsity) but with a high failure rate; Policy B has the lowest expected cost but might result in a situation that costs more to recover; Policy C has a slightly higher expected cost than Policy B but lower variance in cost (risk), which can be considered as a safer Policy.

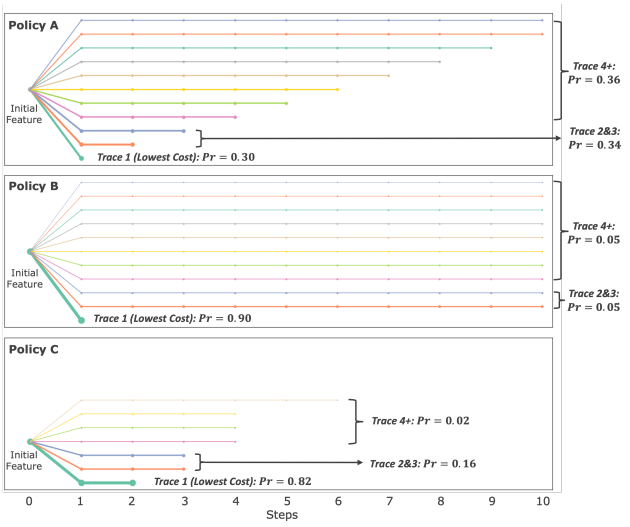

Figure 2: Recourse policy visualization, highlighting variance in cost for three recourse policies from Figure 1; thickness of each outcome (line) is proportional to the probability.

vector of feature changes, recourse can also provide a series of CEs or a sequence of actions [21, 31] that incrementally change users' features and bring them closer to the features with the desired outcome. Some key desiderata to evaluate CEs are [12]: (1) *validity*: whether it gives the desired outcome, (2) *proximity*: how much the changes are, as measured by a distance function, and (3) *sparsity*: how many features are changed, and (4) *realism*: how realistic the recourse recommendations are for an individual, including the feasibility of actions. However, using CEs to find recourse policies does not necessarily result in a sufficiently ***safe*** recourse policy, because they can ignore the risk of taking actions, which may (probabilistically) leave a person in a worse situation. Such a recourse policy may even be dangerous to suggest. For instance, asking a person to change jobs may result in them losing their current job and being jobless (as illustrated in Recourse Policy B in Figure 1). Finding alternatives with lower risks but a slightly higher expected cost may be preferred by an individual. In the context of CE methods, this means that sometimes a more distant state (using a CE measure) of feature values may be a better recourse target if the actions required to reach it carry less risk of higher costs.

To explicitly incorporate risk into algorithmic recourse, our work introduces the problem of computing *safer algorithmic recourse* (**SafeAR**). This has hitherto not been discussed in the literature on algorithmic recourse. The objectives of SafeAR are to suggest different recourse policies with different risk profiles and to empower the affected individual with risk-averse alternatives to decide for themselves.[1] Reinforcement learning (RL) methods can compute such recourse policies. Typically, an RL policy finds the best actions for states, maximizes the expected reward (or minimizes the expected cost), and can incorporate uncertainty in cost and action effects. To account for risk, we incorporate the variance in costs during policy computation and connect risk-sensitive reinforcement learning ideas [3, 7, 16] with algorithmic recourse. Our contributions are:

- Develop the concept of SafeAR, highlighting the value of considering risks in algorithmic recourse, which existing recourse measures do not cover.
- Formulate algorithmic recourse problems as Finite Horizon Markov Decision Processes (MDPs) and demonstrate a method (Greedy Risk-Sensitive Value Iteration, G-RSVI) to compute risk-aware policies for finite horizon MDPs
- Introduce succinct measures of risk into the evaluation of algorithmic recourse, borrowing from the financial literature; these measures are Value at Risk (VaR) [15] and Conditional Value at Risk (CVaR) [34].
- Evaluate the policies with different risk profiles computed by G-RSVI methodology on two real datasets (UCI Adult Income, German Credit) and assess the policies in terms of risk measures, as well as sparsity and proximity to show that the latter do not implicitly factor in risks.
- Conduct an initial investigation into the disparity between gender groups in terms of risks in the aforementioned datasets.

2 MOTIVATING EXAMPLE

To better illustrate the concept of SafeAR with risk-aware policies, consider the following motivating example on loan approvals (Figure 1). A company uses a trained black box ML model to determine loan approvals. The model uses a set of features of the loan applicant (housing, job, savings, age, and education) and initially rejects the applicant. In this recourse example, the action costs are in terms of

[1] SafeAR does not advocate for the policy with the lowest risk as it may have a higher average cost. The emphasis is to provide multiple recourse policies for individuals to choose from, some of which can be safer and more suitable based on their risk tolerance.

discrete time units, each action taken has a probability of success, and failure could transition into a less favorable state. Let us now look at three recourse policies that could be given to the applicant (these are also illustrated in Figure 1):

- *Policy A: Nearest CE, with Expected Cost 3.3*: This could be found by a recourse algorithm that optimizes for feature sparsity. It would require the applicant to *Own-a-House* (one feature change). This policy ignores the uncertainty in the applicant's ability to purchase a house within 1 month (time cost), and there is a 70% chance that the applicant would remain in the same state. So the expected time cost would be much more than 1 month.
- *Policy B: Risk-Neutral Policy with Expected Cost 1.5*: This policy only optimizes the expected cost. It requires the applicant to *Change-Job*, and doing so helps increase the savings and reach the desired outcome with 90% probability. However, there is a small chance (10%) that this action would result in losing their current job and ending up unemployed, from which the cost to recover would be higher. The expected total cost when considering probabilities is still lower than Policy A. If optimizing for expected cost alone, this policy would be returned and has the potential to lead the applicant to a worse situation in which they would incur a high cost to recover from. This is the type of risk in a recourse policy that a user might want to know about and manage.
- *Policy C: Risk-Averse Policy with Expected Cost 2.2*: It provides a safer policy to the applicant, where failures do not lead to a significantly worse situation. The actions for this recourse are *Improve-Education-in-Part-Time* and then *Increase-Savings*. The risk of higher costs in this policy is lower than in Policy B, but it has a higher expected cost. Policy C might not be found by methods that minimize proximity, as improving educational background can be a more distant or larger change in a distance function than getting a higher paying job.

Figure 2 illustrates the probabilities of possible outcome trajectories and their associated costs for this example. This is a visualization paradigm in which we capture the probability of an outcome trajectory by line thickness, and the costs are along the x-axis. With the risk-averse Policy C, the applicant is able to receive the desired outcome in 3 time-steps (cost) with 98% probability, and the risk of it taking more than 3 is much less than Policy A or B, even if the expected cost is more than Policy B. Computing such diverse policies (diverse in terms of risk) and surfacing the risk-information to empower the affected individual is the motivation behind SafeAR.

3 RELATED WORK

Existing algorithmic recourse methods [24] can be grouped into three categories: one set of methods involves finding the nearest CEs as the smallest changes to the individual's feature vector. Solutions focus on *proximity* [46], *sparsity*, and *diversity* [19, 22, 28, 41] using multi-objective optimization [9] and decision trees [20]. Also, generative algorithms [2, 17] are used to ensure *plausibility*, by generating CEs within data distributions. These do not give a sequence of actions or policy to follow and have no mention of risk.

In the second category, recourse is achieved by recommending a sequence of actions [37, 40] or by providing a path over the feature-space along dense regions of the data manifold [31] considering *feasibility* and *actionability*. There are methods that also incorporate *causality* through structural causal models (SCMs) [36] to explicitly model inter-variable causal relationships [26] and provide an ordered sequence of CEs [21, 29]. Lastly, robust recourse methods [30, 39] address the issues for data changes and model parameter shifts. None of these methods consider the risk of higher costs due to the probability of adverse outcomes.

For computing risk-aware recourse policies, we turn to the reinforcement learning (RL) literature. RL methods can provide recourse policies that consider uncertainties in transitions when taking recourse actions. There is existing work that models the recourse problem as MDPs. "ReLAX" [6] generates recourse plans by deep reinforcement learning but under deterministic feature transitions, ignoring uncertainties and thus risk. FASTAR [43] presents a framework that translates an algorithmic recourse problem into a discounted MDP and demonstrates comparable recourse performance as CE methods. Although FASTAR includes uncertainties in the model, it only optimizes for the expected cost and does not incorporate any risk measures. Our method for SafeAR in this work (G-RSVI) computes recourse policies by considering both the expected cost and the risk of higher costs.

One way of measuring risks in cost in RL is through the variance in the total cost (over all steps) [32, 38]. There are also other RL methods that can factor in risks in policies [5, 8, 11, 16]. To communicate the idea of SafeAR in this work, we use a modification of value iteration (G-RSVI) to incorporate the cost-variance trade-off into the computation to get risk-averse policies. MDPs can naturally incorporate action costs, probabilistic action dynamics, action feasibility, and causal constraints. These can all be personalized to the recipient, as properties like action costs can be unique to each person. To the best of our knowledge, SafeAR is the first attempt to connect the literature on *risk-sensitive* RL to algorithmic recourse.

4 SAFE ALGORITHMIC RECOURSE

4.1 Algorithmic Recourse Problem Statement

Let $f : \mathcal{X} \to \mathcal{Y}$ be a decision function operationalized by an ML algorithm or model, where $x \in \mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_D$ is the set of instances described by D features of an individual, and $\mathcal{Y} = \{y^-, y^+\}$ are the unfavorable and favorable decision outcomes, respectively. An individual with features x_o initially gets an unfavorable outcome $f(x_o) = y^-$, and the general objective of algorithmic recourse is to find actions resulting in a path $x_o, \ldots, x^*$ that leads to final feature instance x^* so that $f(x^*) = y^+$. Our work is agnostic to the type of ML model f and only requires the model outputs to be categorized into unfavorable outcomes y^- and favorable outcomes y^+. For simplicity of discourse, we use a binary classifier for f.

4.2 Risk-Aware Recourse Policies Using Finite Horizon Markov Decision Processes

To compute SafeAR recommendations, we frame the problem as solving a finite horizon Markov Decision Process (MDP), defined as a tuple of $\langle S, A, T, R, H \rangle$. H is the maximum number of steps in the

finite horizon MDP and policies, and $h := [1 : H]$ is the step number in the horizon.

States (S). S is a set of all possible states for individuals in the recourse. Each of the states maps to one instance ($x \in \mathcal{X}$) in the combined feature space (input space) of the decision model f. For a valid state space, there must exist a mapping $g := S \rightarrow X$, where $\forall s \in S, \exists x \in X$ such that $g(s) = x$. In this work, we keep the mapping g as one-to-one, meaning the state space is equivalent to the feature space. However, the state space S can be *richer* than the feature space $\mathcal{X}$ because the states and actions for recourse can involve more or different features than the ones used in the decision model f. For example, "food expenses" can be a feature in the recourse state space S associated with actions for saving money, but not in the feature space f for a loan-approvals model. Using only the same features as in f can be insufficient for computing recourse policies, as those features may not cover the states and actions that a person actually affects during the recourse. This gives us a reason to expect a separate model for action transitions and action costs for recourse, rather than assuming it can be extracted from the data used in f. Another strong reason to expect a separate action model is for recourse personalization, as advocated for in [42].

Actions (A) and Transitions (T). For the state space S, we have a set of feasible actions $a \in A$. The effect of an action can change multiple features. The features can be categorized into three types [23]: 1) immutable features (e.g., birthplace), 2) mutable and actionable features (e.g., occupation, bank balance) that define the action space of the recourse, 3) mutable but non-actionable features (e.g., credit score) that cannot be directly modified by an individual. Mutable features can be modified as a consequence of changing other features. The state transition model would need to capture causal relationships between features and ensure the realism of recourse. The state transition model $T := p(s'|s, a)$ is defined as the transition probability between two states ($\{s, s'\} \in S \times S$) given the action a.

Rewards (R). $R := r(s, a, s'; f)$ is the reward or cost incurred by reaching state s' by performing an action a at state s. "Reward" and $r(.)$ are the typical terms and notations used in RL literature, but rewards can be positive or negative (cost). We will henceforth use "cost" in this work since we are focusing on the recourse cost to the recipient, i.e. $r(s, a, s'; f)$ is the cost incurred to the recipient when the transition (s, a, s') occurs during the recourse. Additionally, the ML model f indicates which states provide the favorable outcome ($f(s) = y^+$). In these states, no more actions are needed in the recourse. To capture this, we add a zero-cost action in all favorable (goal) states and they result in a transition to the same state.

As for the real-world semantics of the cost, the cost can be a measure of one or more cost-factors such as elapsed time, material expenses, opportunity cost, etc. Additionally, the cost maybe averaged across a group or tailored for each person, which requires domain knowledge; domain knowledge is also needed for defining feasible actions and transitions. CE methods such as DiCE [28] and FACE [31] also require domain knowledge to design the distance function (which captures the cost to humans), which can be in terms of how much the state changes by an action using *sparsity* of feature changes, or *proximity* of the recipient's state changes over some feature-distance functions.

Algorithm 1 G-RSVI

Input: recourse MDP$\langle S, A, T, R, H\rangle$, ML model f
Parameters: risk aversion level $\beta \in [0, \infty]$

$V_{H+1}(s) \leftarrow 0, \forall s \in S$
for step $h = H, H-1, \ldots, 1$ **do**
  for each state $s \in S$ **do**
    for each action $a \in A$ **do**
      $r(s') \leftarrow$ get reward $R(s, a, s'; f)$
      $p(s') \leftarrow$ get transition probability $T(s'|s, a)$
      $\mu \leftarrow \sum_{s'} p(s')[r(s') + V_{h+1}(s')]$
      $\sigma^2 \leftarrow \sum_{s'} p(s')[r(s') + V_{h+1}(s') - \mu]^2$
      $Q_h(s, a) \leftarrow \mu - \beta\sigma$ (Equation 3)
    end for
    $V_h(s) \leftarrow \max Q_h(s, \cdot)$
    $\pi_h(s) \leftarrow \text{argmax}\, Q_h(s, \cdot)$
  end for
end for
return recourse policy $\pi_h(s)$

Recourse Policies. A recourse policy is the same as an MDP policy, expressing how to act in each state of each step in the horizon to get to a favorable state. This is formalized as $\pi = (\pi_1 \ldots \pi_H)$, where $\pi_i := S \rightarrow A$ maps each state to an action for each step i.

5 SAFEAR METHODOLOGY

In this section, we present a method for computing risk-averse recourse policies and measures to evaluate the risk for SafeAR.

5.1 Greedy Risk-Sensitive Value Iteration

In this section, we present a greedy algorithm to compute risk-averse policies by incorporating cost-variance into the policy computation. We first define $\hat{R}_h^\pi(s) = \sum_{i=h}^{H} r(s_i, a_i, s_{i+1})|(s_h = s), \pi$ as the total cost accrued over the horizon H from a rollout obtained by following a policy π starting at state s and step h. In risk-neural settings, the recourse policy π maximizes the expected total cost $\mathbb{E}[\hat{R}_h^\pi(s)]$ or the mean value $\mu[\hat{R}_h^\pi(s)]$. G-RSVI additionally considers the variance in cost to manage risk and seeks to find a policy π to maximize the following value function:

$$V_1^\pi(s) = \mu(\hat{R}_1^\pi(s)) - \beta \cdot \sigma(\hat{R}_1^\pi(s)), \tag{1}$$

for each state s starting at the first step $h = 1$. We denote $V_h^\pi(s)$ as the risk-sensitive value of state s in step h by following policy π. $\beta \geq 0$ is the tuning parameter that represents each individual's risk profile, and a higher value means more risk averse. When $\beta = 0$, the problem reduces to finding the policy with the least expected cost only, which is the standard optimization objective in MDPs. Here, σ returns the standard deviation of the total cost, and σ^2 returns the variance. In G-RSVI, we optimize Equation 1 by greedily maximizing $V_h^\pi(s)$ at each step starting from the last step H and moving backward to the first step. At each step h, the action is selected to maximize the risk-sensitive value using:

$$V_h = \max_a \mu[r(\cdot) + V_{h+1}] - \beta\sigma[r(\cdot) + V_{h+1}], \tag{2}$$

where V_{h+1} represents the values computed for the $h+1$ step, and 0 when $h = H$. The risk-sensitive action value or Q-value $Q_h(s, a)$ at

step h is defined as:

$$Q_h(s,a) = \mathbb{E}_{s'}[r(s,a,s') + V_{h+1}(s')] - \beta\sigma[r(s,a,s') + V_{h+1}(s')]. \quad (3)$$

If only optimizing the expected reward ($\beta = 0$), this procedure would find the optimal policy because the optimal sub-structure assumption for dynamic programming holds. However, G-RSVI is not guaranteed to find the optimal policy for Equation 1. It does, however, provide one straightforward way to incorporate risks into recourse policy computation, and it completes computation with a single sweep over the state and horizon space, making it a fast approach. There are a variety of heuristic methods one can use with different trade-offs to compute risk-aware policies. In this work, to focus on the exposition of the concept of SafeAR, we limit our approach to discrete states, discrete actions, and finite horizon MDPs.

Our G-RSVI algorithm is shown in Algorithm 1. We compute the policy by sweeping backward from the last horizon step (Line 2). For all state-action pairs at each step, the action values $Q_h(s,a)$ are computed by Equation 3 (Lines 5-9). The best action for each state in each step is then chosen by the one with maximal $Q_h(s,a)$. It also gives us the state value $V_h(s)$ and the policy for each step (lines 11, 12). Other ways of scoring values and selecting actions can be used other than Equation 3 in our algorithm. For example, one can optimize for CVaR, although that requires specifying a confidence level. For this initial work on SafeAR with risk-aware policies, we limit the scope to using Equation 3. This also helps us compare against FASTAR [43] as the baseline, where the FASTAR equivalent policy is obtained by setting $\beta = 0$ since FASTAR optimizes for the expected value only. We leave the analysis of different risk-sensitive algorithms for recourse to future work.

5.2 Risk Measures for Recourse Policies

To evaluate the risk associated to a recourse policy, we propose the following measures.

Success Rate (ρ_H): It estimates the probability of success within the finite horizon H by following the recourse policy. For example, $\rho_5 = 0.9$ means a favorable outcome state will be reached within 5 steps 90% of the time. This is not equivalent to *validity*, which only determines whether an instance of feature combinations with a favorable decision exists. ρ_H is affected by the uncertainty of action outcomes in the recourse policy.

Mean-Variance Cost ($\mu_{cost}, \sigma^2_{cost}$): It computes the expected value and variance in the total cost of following recourse policies. Since the distribution of costs is not necessarily Gaussian, these statistics can be misleading or hard to interpret. Hence, we propose additional measures.

Value at Risk (VaR_α): VaR [15] is to provide a succinct probabilistic guarantee on the recourse policy cost. We evaluate VaR [15] of the recourse cost to answer the question "What is the highest cost at a given level of cumulative probability (confidence level)". For example, $\text{VaR}_{95} = 5.6$ means that with 95% probability, the recourse cost is at most 5.6. Formally, assuming the total cost of recourse x_c is the value of a random variable X_c with a cumulative probability distribution $F_X(x_c)$, under confidence level $\alpha \in [0,1]$ VaR_α is computed as:

$$\text{VaR}_\alpha(X_c) = \min\{x_c | F_X(x_c) \geq \alpha\}. \quad (4)$$

Conditional Value at Risk (CVaR_α): CVar [34] is a complementary measure to VaR and tells us the expected worst case cost when the cost exceeds the threshold given by VaR_α value. For example, $\text{CVaR}_{95} = 8.4$ means that when the cost exceeds the 95th percentile cost, the average cost for those cases is 8.4. It is computed as:

$$\text{CVaR}_\alpha = \mathbb{E}[x_c | x_c > \text{VaR}_\alpha]. \quad (5)$$

6 EXPERIMENTAL RESULTS

Motivated by the datasets used in the algorithmic recourse literature, we evaluate our method on the following two datasets: Adult Income Dataset (AID) (32561 data points) [4] and German Credit Dataset (GCD) (1000 data points) [14] and show how risk measures vary with different recourse policies. In AID, the recourse is to help individuals earn an income greater than 50,000. In GCD, the recourse is to help get a loan approval by reaching a good credit standing. Here we consider the version of GCD [18] with 9 features. To process the datasets for G-RSVI, we convert continuous feature values into discrete values (details included in Appendix A.1[2]). We then train random forest classifiers for both datasets. Dataset features, feature state dimensions, and classifier accuracies are reported in Table 2.

Transitions and Rewards. Since the datasets for recourse do not have associated action models, we use qualitative assumptions (domain knowledge) on relative differences in action costs and success likelihood to define the action costs $r(\cdot)$ and transition model $p(\cdot)$. Similar to FASTAR [43], we assume "*improve-education*" or "*improve-skill*" actions would lead to an age increase as causal constraints, and we treat "Age" as a mutable but non-actionable feature. These two actions require more time and effort, and therefore the action cost would be larger than other actions such as "*increase-work-hours*". The transition probabilities are heuristically set by domain knowledge. For example, the probability of earning a Ph.D. degree is lower than earning a Bachelor's. For the details on the model values, we refer the reader to Appendix A.2 (see footnote 2) for an exhaustive list of model transition probabilities and costs. Additional results from a different model using the same qualitative assumptions are also provided in Appendix A.5 to show the G-RSVI method and results are not specific to a single model.

Baselines. To our knowledge, our work is the first to address risks in algorithmic recourse. Among the existing recourse approaches, only FASTAR [43] formulates recourse problems as MDPs *and* allows for stochastic transitions. FASTAR sets rewards in terms of distance measures between states. No matter what reward function is used—either distance-based or user-defined cost—and how transition probabilities are defined—either extracted from a dataset or tuned domain knowledge—FASTAR only seeks to find the recourse policy that maximizes the expected total rewards (risk-neural). This is what a standard algorithm for MDP (value or policy iteration) would find. In our experiments, the policy that maximizes expected total reward corresponds to the risk-neutral policy ($\beta = 0$), and this is the baseline which risk-averse policies compare against. We select

[2]All supplemental materials are available and accessible *through this link*

Dataset	Policy	$\rho_{H=12}$	$(\mu_{cost}, \sigma^2_{cost})$	VaR_{80}	CVaR_{80}	VaR_{95}	CVaR_{95}	Spars.	Proxi.
Adult Income ($n = 25923$)	$\beta = 0$	0.994	(**3.49**, 1.23)	3.81	6.31	4.76	7.53	**2.09**	**2.87**
	$\beta = 0.25$	0.994	(3.51, 0.89)	**3.64**	**6.10**	4.46	7.43	2.16	3.06
	$\beta = 0.5$	0.993	(3.59, **0.77**)	3.66	6.11	**4.44**	**7.40**	2.21	3.18
(Example Instance)	$\beta = 0$	1.000	(**4.63**, 1.86)	5.80	8.54	6.80	9.80	3.92	**3.92**
	$\beta = 0.5$	1.000	(4.79, **0.13**)	**4.80**	**6.80**	**4.80**	**6.80**	**3.00**	4.86
	$\beta = 0.75$	1.000	(4.79, **0.13**)	**4.80**	**6.80**	**4.80**	**6.80**	**3.00**	4.86
German Credit ($n = 281$)	$\beta = 0$	1.000	(**1.65**, 0.48)	1.96	3.66	2.63	4.56	**1.26**	**1.33**
	$\beta = 0.25$	1.000	(1.67, 0.35)	**1.87**	**3.51**	2.51	4.50	1.34	1.43
	$\beta = 0.5$	1.000	(1.70, **0.30**)	1.90	3.56	**2.48**	**4.40**	1.40	1.50
(Example Instance)	$\beta = 0$	1.000	(**2.48**, 3.49)	**4.00**	6.13	7.00	8.33	**1.00**	**1.00**
	$\beta = 0.5$	1.000	(2.81, 1.19)	**4.00**	**5.44**	**5.00**	**6.00**	**1.00**	2.00
	$\beta = 0.75$	1.000	(3.87, **0.65**)	4.40	5.50	5.40	6.67	2.00	3.00

Table 1: Evaluating recourse policies with different risk-aversion levels β and horizon $H = 12$ for AID, GCD, and two selected example instances. n denotes the number of instances with unfavorable outcomes used for evaluation. For each policy, we first run 100 trails for each instance in the dataset and measure the metrics among the valid recourse trials. Then, the average across all instances for each metric is computed. The best metric values among the policies are highlighted in bold.

Dataset (#Features)	Immutable Features	#States	ML Model (Accuracy)
AdultIncome (8)	Gender, Race, Martial Status	57600	Rand.Forest (0.81)
GermanCredit (9)	Sex, Purpose, Credit Amount	147456	Rand.Forest (0.76)

Table 2: Dataset Overview

$\beta = 0.25, 0.50, 0.75$ for generating risk-averse recourse policies, and higher β indicates higher risk-aversion.

Performance Evaluation. Table 1 reports the risk measures for each experimental setting. The horizon is set to 12. We also present sparsity (measured by L_0 distance) and proximity (measured by L_0 distance for nominal features + L_1 distance for ordinal and numerical features) between initial and final states. All measures are averaged over the entire dataset, as well as the measures for two example instances (sample from $\mathcal{X}$). Recourse policies are computed using the same cost and transition functions for all instances in the dataset. In the results, we see that for both datasets, more risk-averse policies (higher β values) can provide recourse with less variance in cost σ^2_{cost} but often require a higher mean cost μ_{cost}. For the same α confidence level, risk-averse policies give lower costs in VaR and CVaR than risk-neutral policies. Also, for the example instance in GCD dataset, variance in cost $\sigma^2 = 0.65$ at risk-aversion level $\beta = 0.75$ is significantly lower than the $\sigma^2 = 3.49$ at $\beta = 0$. For the example instance in AID dataset, increasing risk-aversion to $\beta = 0.75$ would not find a different policy than $\beta = 0.5$, which can happen if the same relative ordering of state values $V_h(s)$ is found at each step. In the results, we observe that low sparsity and proximity do not correspond to risk-averse policies, meaning optimizing for them would not necessarily factor in risks.

Visualizing Risks in Recourse Policies. In Figure 3, we use our policy-risk visualization for a set of policies from the GCD dataset. For each policy, we show the most probable outcomes (rollouts), the length of each trace corresponds to the total cost, and the thickness

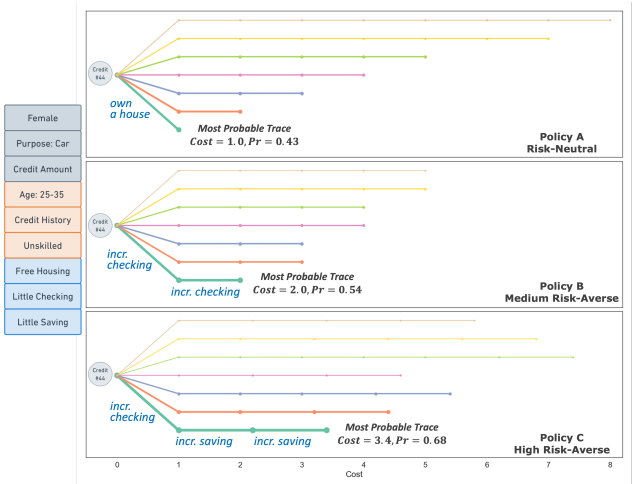

Figure 3: Policy visualization for an Example Instance in German Credit

of each trace corresponds to the probability of the outcome. This approach visualizes the variability in cost, which can help a person get an intuition of their risk in addition to the recommended actions.

Exploring Risks across Gender. Inspired by prior work that investigated disparity in recourse between different groups [13, 33, 35, 45], we now look at the disparity that may exist in risk measures across two gender groups (male and female) provided in AID and GCD at the same risk-aversion level. Table 3 reports the same risk measures (averaged) for female and male groups in both datasets. Information on the p-values for statistical significance of the difference between the groups is provided in Appendix A.4 (see footnote 2). We define disparity in VaR between the two gender groups by following the same recourse policy computed with a risk-aversion level β as: $\Delta\text{VaR}^{\beta}_{95} = |\text{VaR}^{\text{Female}}_{95} - \text{VaR}^{\text{Male}}_{95}|$. The disparity for other measures is similarly computed.

All measures are in favor of the male group (highlighted in green shades) for both datasets, meaning that for the policy with the same risk-aversion level given to females and males, we expect females

Dataset	Policy	$\rho_{H=12}$	$(\mu_{cost}, \sigma^2_{cost})$	VaR_{80}	CVaR_{80}	VaR_{95}	CVaR_{95}	Spars.	Proxi.
Adult Income (Female, n = 9824)	$\beta = 0$	0.988	(**4.56**, 1.41)	4.76	7.07	5.70	8.00	2.58	**3.77**
	$\beta = 0.25$	0.987	(4.57, 1.18)	**4.64**	**6.95**	**5.48**	**7.84**	**2.55**	3.87
	$\beta = 0.5$	0.985	(4.61, **1.11**)	4.66	6.99	5.51	7.89	**2.55**	3.93
(Male, n = 16099)	$\beta = 0$	0.997	(**2.84**, 1.12)	3.27	5.79	4.30	7.26	**1.79**	**2.32**
	$\beta = 0.25$	0.997	(2.87, 0.72)	**3.08**	5.51	3.95	7.15	1.93	2.56
	$\beta = 0.5$	0.998	(2.98, **0.57**)	3.10	**5.48**	**3.92**	**7.08**	2.01	2.73
German Credit (Female, n = 103)	$\beta = 0$	1.000	(**1.72**, 0.56)	2.08	3.70	2.84	4.66	**1.25**	**1.36**
	$\beta = 0.25$	1.000	(1.75, 0.38)	**2.00**	**3.61**	2.67	4.63	1.33	1.47
	$\beta = 0.5$	1.000	(1.79, **0.33**)	2.02	3.64	**2.59**	**4.51**	1.41	1.57
(Male, n = 178)	$\beta = 0$	1.000	(**1.61**, 0.43)	1.89	3.64	2.61	4.62	**1.27**	**1.32**
	$\beta = 0.25$	1.000	(1.62, 0.37)	**1.81**	**3.50**	2.41	4.42	1.35	1.41
	$\beta = 0.5$	1.000	(1.65, **0.29**)	1.84	**3.50**	**2.33**	**4.29**	1.40	1.46

Table 3: Evaluating recourse policies across gender for Adult Income and German Credit datasets; the same evaluation procedures are followed as Table 1. Among risk measures, the cells shaded in green indicate that the corresponding gender group is exposed to less risk under the policy with the same risk-aversion level.

would get higher variance in cost (σ^2_{cost}), higher cost at the VaR confidence level of $\alpha = 80$ and $\alpha = 95$, and higher costs in the expected worst case scenarios (CVaR) for those confidence levels. In AID, we also noticed that when increasing the risk-aversion, the disparity of risk measures between two groups becomes *larger*. We observe $\Delta\text{VaR}^{\beta}_{95}$ increases from 1.4 to 1.59 as β increases, and similar trends are observed for the difference in σ^2_{cost} and CVaR in AID. This trend indicates that the more we want to achieve risk-aversion, the greater the disparity in risk exposure between the two gender groups in AID. However, in GCD, the difference in σ^2_{cost} and $\Delta\text{VaR}^{\beta}_{80}$ do not consistently increase with increased risk-aversion. The disparity between males and females across all measures of risk still exists in GCD. We recall that the same action costs and transitions are used for both males and females and were not informed by gender. We also present these results for a different action model in Appendix A.5 (see footnote 2 for the link to Appendix). The only difference is the decisions made by the model f for different groups, which affects the number of steps or cost to reach the favorable state ($f(x) = y^+$). It can be worthwhile for recourse providers to test for and consider such disparities, and it motivates further discussion on risk disparity in algorithmic recourse.

7 DISCUSSION AND CONCLUSIONS

The motivation behind Safer Algorithmic Recourse (SafeAR) is to offer recourse policies with different risk profiles. This enables affected individuals to be more aware of the risks and to help them make informed decisions based on their risk tolerance. We connect ideas from risk-sensitive reinforcement learning with the algorithmic recourse literature and propose an algorithm G-RSVI that can provide risk-averse recourse policies for individuals with different risk profiles. In our experiments with the AID and GCD datasets, we showed that the recourse policies generated by G-RSVI were better in terms of the risk measures as compared to the existing risk-neutral approaches. The policy risk was evaluated through cost-variance, VaR, CVaR, and success rate. In addition, we observed that policies with better sparsity and proximity scores need not correspond to risk-averse policies. Lastly, in our experiments, we observed discrepancies between gender groups in risk measures for the same risk-aversion setting and action model; this motivates further studies on recourse fairness in terms of risk exposure.

REFERENCES

[1] Gah-Yi Ban, Noureddine El Karoui, and Andrew EB Lim. 2018. Machine learning and portfolio optimization. *Management Science* 64, 3 (2018), 1136–1154.

[2] Alejandro Barredo-Arrieta and Javier Del Ser. 2020. Plausible Counterfactuals: Auditing Deep Learning Classifiers with Realistic Adversarial Examples. *International Joint Conference on Neural Networks (IJCNN)* (2020), 1–7. https://doi.org/10.1109/IJCNN48605.2020.9206728

[3] Nicole Bäuerle and Ulrich Rieder. 2014. More risk-sensitive Markov decision processes. *Mathematics of Operations Research* 39, 1 (2014), 105–120.

[4] Barry Becker and Ronny Kohavi. 1996. Adult. UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C5XW20.

[5] Vivek S Borkar. 2010. Learning algorithms for risk-sensitive control. *Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems–MTNS* 5, 9 (2010).

[6] Ziheng Chen, Fabrizio Silvestri, Jia Wang, He Zhu, Hongshik Ahn, and Gabriele Tolomei. 2022. ReLAX: Reinforcement Learning Agent Explainer for Arbitrary Predictive Models. *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM '22)* (2022), 252–261. https://doi.org/10.1145/3511808.3557429

[7] Yinlam Chow, Aviv Tamar, Shie Mannor, and Marco Pavone. 2015. Risk-sensitive and robust decision-making: a cvar optimization approach. *Advances in neural information processing systems* 28 (2015).

[8] Yinlam Chow, Aviv Tamar, Shie Mannor, and Marco Pavone. 2015. Risk-sensitive and robust decision-making: a cvar optimization approach. *Advances in neural information processing systems* 28 (2015).

[9] Susanne Dandl, Christoph Molnar, Martin Binder, and Bernd Bischl. 2020. Multi-Objective Counterfactual Explanations. *Parallel Problem Solving from Nature – PPSN XVI* (2020), 448–469.

[10] Anirban Datta. 2016. US Health Insurance Dataset - Kaggle. https://www.kaggle.com/datasets/teertha/ushealthinsurancedataset. note="Accessed: 2023-08-14.

[11] Yingjie Fei, Zhuoran Yang, and Zhaoran Wang. 2021. Risk-Sensitive Reinforcement Learning with Function Approximation: A Debiasing Approach. *Proceedings of the 38th International Conference on Machine Learning* 139 (18–24 Jul 2021), 3198–3207.

[12] Riccardo Guidotti. 2022. Counterfactual explanations and how to find them: literature review and benchmarking. *Data Mining and Knowledge Discovery* (2022). https://doi.org/10.1007/s10618-022-00831-6

[13] Aparajita Haldar, Teddy Cunningham, and Hakan Ferhatosmanoglu. 2022. RAGUEL: Recourse-Aware Group Unfairness Elimination. *Proceedings of the 31st ACM International Conference on Information and Knowledge Management* (2022), 666–675. https://doi.org/10.1145/3511808.3557424

[14] Hans Hofmann. 1994. Statlog (German Credit Data). UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C5NC77.

[15] Glyn A Holton. 2013. *Value-at-Risk: Theory and Practice, Second Edition*. Online. https://www.value-at-risk.net

[16] Ronald A Howard and James E Matheson. 1972. Risk-sensitive Markov decision processes. *Management science* 18, 7 (1972), 356–369.

[17] Shalmali Joshi, Oluwasanmi Koyejo, Warut Vijitbenjaronk, Been Kim, and Joydeep Ghosh. 2019. Towards Realistic Individual Recourse and Actionable Explanations in Black-Box Decision Making Systems. *arXiv* (2019). https://doi.org/arXiv:

1907.09615

[18] Kaggle. 2016. German Credit Risk - UCI MACHINE LEARNING. https://www.kaggle.com/datasets/uciml/german-credit. note="Accessed: 2023-06-18.

[19] Kentaro Kanamori, Takuya Takagi, Ken Kobayashi, and Hiroki Arimura. 2020. DACE: Distribution-Aware Counterfactual Explanation by Mixed-Integer Linear Optimization. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence* (7 2020), 2855–2862.

[20] Kentaro Kanamori, Takuya Takagi, Ken Kobayashi, and Yuichi Ike. 2022. Counterfactual Explanation Trees: Transparent and Consistent Actionable Recourse with Decision Trees. *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics* 151 (2022), 1846–1870.

[21] Kentaro Kanamori, Takuya Takagi, Ken Kobayashi, Yuichi Ike, Kento Uemura, and Hiroki Arimura. 2021. Ordered Counterfactual Explanation by Mixed-Integer Linear Optimization. *Proceedings of the AAAI Conference on Artificial Intelligence* 35, 13 (2021), 11564–11574. https://doi.org/10.1609/aaai.v35i13.17376

[22] Amir-Hossein Karimi, Gilles Barthe, Borja Balle, and Isabel Valera. 2020. Model-Agnostic Counterfactual Explanations for Consequential Decisions. *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics* 108 (26–28 Aug 2020), 895–905.

[23] Amir-Hossein Karimi, Gilles Barthe, Bernhard Schölkopf, and Isabel Valera. 2022. A Survey of Algorithmic Recourse: Contrastive Explanations and Consequential Recommendations. *Comput. Surveys* 55, 5 (2022), 1–29. https://doi.org/10.1145/3527848

[24] Amir-Hossein Karimi, Bernhard Schölkopf, and Isabel Valera. 2021. Algorithmic Recourse: From Counterfactual Explanations to Interventions. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (2021), 353–362. https://doi.org/10.1145/3442188.3445899

[25] Hua Li, Yumeng Cao, Siwen Li, Jianbin Zhao, and Yutong Sun. 2020. XGBoost model and its application to personal credit evaluation. *IEEE Intelligent Systems* 35, 3 (2020), 52–61.

[26] Divyat Mahajan, Chenhao Tan, and Amit Sharma. 2019. Preserving Causal Constraints in Counterfactual Explanations for Machine Learning Classifiers. *arXiv:1912.03277* (2019).

[27] Patrick E McKnight and Julius Najab. 2010. Mann-Whitney U Test. *The Corsini encyclopedia of psychology* (2010), 1–1.

[28] Ramaravind K. Mothilal, Amit Sharma, and Chenhao Tan. 2020. Explaining Machine Learning Classifiers through Diverse Counterfactual Explanations. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (2020), 607–617. https://doi.org/10.1145/3351095.3372850

[29] Philip Naumann and Eirini Ntoutsi. 2021. Consequence-Aware Sequential Counterfactual Generation. *Machine Learning and Knowledge Discovery in Databases. Research Track* 12976 (2021), 682–698. https://doi.org/10.1007/978-3-030-86520-7_42

[30] Duy Nguyen, Ngoc Bui, and Viet Anh Nguyen. 2023. Distributionally Robust Recourse Action. *The Eleventh International Conference on Learning Representations* (2023). https://openreview.net/forum?id=E3ip6qBLF7

[31] Rafael Poyiadzi, Kacper Sokol, Raul Santos-Rodriguez, Tijl De Bie, and Peter Flach. 2020. FACE: Feasible and Actionable Counterfactual Explanations. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (2020), 344–350. https://doi.org/10.1145/3375627.3375850

[32] LA Prashanth and Mohammad Ghavamzadeh. 2016. Variance-constrained actor-critic algorithms for discounted and average reward MDPs. *Machine Learning* 105 (2016), 367–417.

[33] Francesca E. D. Raimondi, Andrew R. Lawrence, and Hana Chockler. 2022. Equality of Effort via Algorithmic Recourse. arXiv:2211.11892 [stat.ML]

[34] R. Tyrrell Rockafellar and Stanislav Uryasev. 2000. Optimization of conditional value-at risk. *Journal of Risk* 2, 3 (2000), 21–41. https://doi.org/10.21314/JOR.2000.038

[35] Shubham Sharma, Jette Henderson, and Joydeep Ghosh. 2020. CERTIFAI: A Common Framework to Provide Explanations and Analyse the Fairness and Robustness of Black-Box Models. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (2020), 166–172. https://doi.org/10.1145/3375627.3375812

[36] Shohei Shimizu, Takanori Inazumi, Yasuhiro Sogawa, Aapo Hyvärinen, Yoshinobu Kawahara, Takashi Washio, Patrik O. Hoyer, and Kenneth Bollen. 2011. DirectLiNGAM: A Direct Method for Learning a Linear Non-Gaussian Structural Equation Model. *Journal of Machine Learning Research* 12, 33 (2011), 1225–1248. http://jmlr.org/papers/v12/shimizu11a.html

[37] Ronal Singh, Tim Miller, Henrietta Lyons, Liz Sonenberg, Eduardo Velloso, Frank Vetere, Piers Howe, and Paul Dourish. 2023. Directive Explanations for Actionable Explainability in Machine Learning Applications. *ACM Transactions on Interactive Intelligent Systems* 34 (2023). https://doi.org/10.1145/3579363

[38] Matthew J Sobel. 1994. Mean-variance tradeoffs in an undiscounted MDP. *Operations Research* 42, 1 (1994), 175–183.

[39] Sohini Upadhyay, Shalmali Joshi, and Himabindu Lakkaraju. 2021. Towards Robust and Reliable Algorithmic Recourse. *Advances in Neural Information Processing Systems* (2021). https://openreview.net/forum?id=AuVKs6JmBtY

[40] Berk Ustun, Alexander Spangher, and Yang Liu. 2019. Actionable Recourse in Linear Classification. *Proceedings of the Conference on Fairness, Accountability, and Transparency* (2019), 10–19. https://doi.org/10.1145/3287560.3287566

[41] Arnaud Van Looveren and Janis Klaise. 2021. Interpretable Counterfactual Explanations Guided by Prototypes. *Machine Learning and Knowledge Discovery in Databases. Research Track* (2021), 650–665.

[42] Suresh Venkatasubramanian and Mark Alfano. 2020. The philosophical basis of algorithmic recourse. *Proceedings of the 2020 conference on fairness, accountability, and transparency* (2020), 284–293.

[43] Sahil Verma, Keegan Hines, and John P. Dickerson. 2022. Amortized Generation of Sequential Algorithmic Recourses for Black-Box Models. *Proceedings of the AAAI Conference on Artificial Intelligence* 36, 8 (2022), 8512–8519. https://doi.org/10.1609/aaai.v36i8.20828

[44] Julius Von Kügelgen, Amir-Hossein Karimi, Umang Bhatt, Isabel Valera, Adrian Weller, and Bernhard Schölkopf. 2022. On the fairness of causal algorithmic recourse. *Proceedings of the AAAI conference on artificial intelligence* 36, 9 (2022), 9584–9594.

[45] J. von Kügelgen, A.-H. Karimi, U. Bhatt, I. Valera, A. Weller, and B. Schölkopf. 2022. On the Fairness of Causal Algorithmic Recourse. *Proceedings of the 36th AAAI Conference on Artificial Intelligence* 9 (2022), 9584–9594. https://doi.org/10.1609/aaai.v36i9.21192

[46] Sandra Wachter, Brent Mittelstadt, and Chris Russell. 2017. Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harv. JL & Tech.* 31 (2017), 841.

[47] Haojun Zhu. 2016. Predicting Earning Potential using the Adult Dataset. https://rpubs.com/H_Zhu/235617