SafeAR: Safe Algorithmic Recourse by Risk-Aware Policies

Haochen Wu[1], **Shubham Sharma**[2], **Sunandita Patra**[2], **Sriram Gopalakrishnan**[2]

[1] University of Michigan, Ann Arbor
[2] J.P. Morgan AI Research
haochenw@umich.edu, shubham.x2.sharma@jpmchase.com, sunandita.patra@jpmchase.com, sriram.gopalakrishnan@jpmchase.com

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 receiving 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 Safe 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 datasets and compare policies with different risk-aversion levels using risk measures and recourse desiderata (sparsity and proximity).

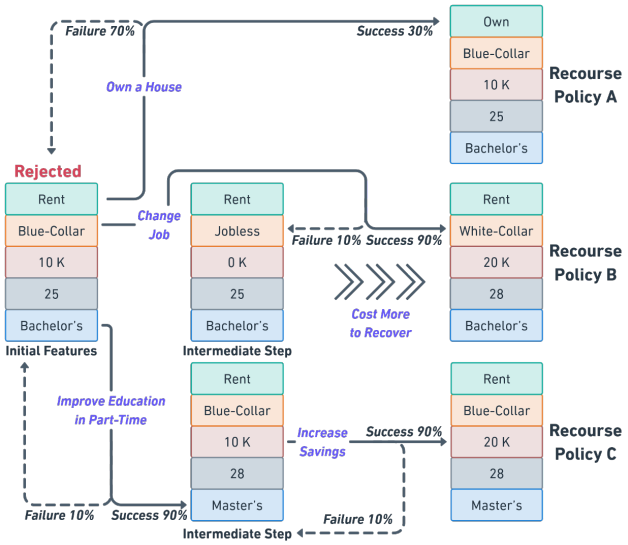

Figure 1: Recourse policies for loan approvals. Policy A has only one feature change (low sparsity) but with a high failure rate; Policy B has the lowest expected cost but might result in a situation that costs more to recover; Policy C has a slightly higher expected cost than Policy B but lower variance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

Introduction

Machine learning (ML) models are increasingly being used to make decisions in a wide array of scenarios including healthcare (Beam and Kohane 2018), insurance premiums (ul Hassan et al. 2021), and loan approvals (Li et al. 2020). Given the impact of ML models on society, the importance of algorithmic recourse has increased (Venkatasubramanian and Alfano 2020). Algorithmic recourse refers to a computed recommendation provided to an end user that 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 (Von Kügelgen et al. 2022; Verma, Hines, and Dickerson 2022; Poyiadzi et al. 2020). 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* here 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, that result in the desired prediction from the model" (Karimi, Schölkopf, and Valera 2021). 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 vector of feature changes, recourse can also provide a series of CEs or a sequence of actions (Poyiadzi et al. 2020; Kanamori et al. 2021) that incrementally change users' features to ultimately achieve the desired outcome. Some key desiderata to evaluate CEs are (Guidotti 2022): (1) *validity*: whether it gives the desired outcome, (2) *proximity*: how much the changes are measured by a distance function, (3) *sparsity*: how many features are changed, and (4) *realism*: how realistic 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 might 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 (set 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 *safe 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 be used to compute such recourse policies. Typically, a policy in RL finds the best action given a (feature) state, which maximizes the expected reward (or minimizes the 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 (Howard and Matheson 1972; Chow et al. 2015a; Bäuerle and Rieder 2014) with algorithmic recourse. Our contributions are:

- Develop the concept of SafeAR by highlighting the value of considering risks in algorithmic recourse, which existing recourse measures do not cover.

[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.

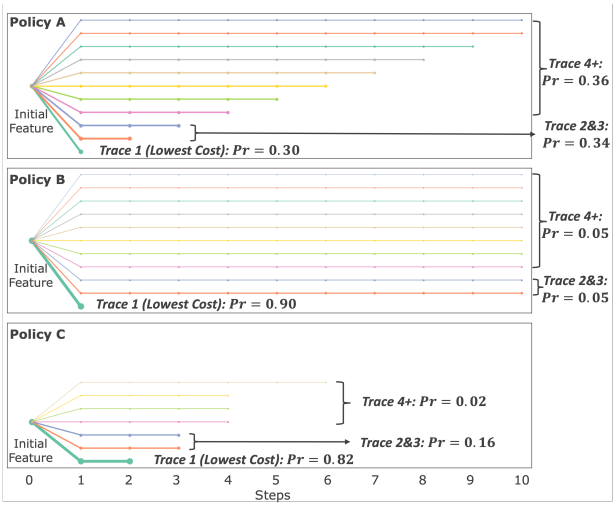

Figure 2: Visualizing risks in three recourse policies from Figure 1. The x-axis indicates the cost, and the line (outcome) thickness indicates the outcome probability.

- Formulate algorithmic recourse problems as Finite Horizon Markov Decision Processes (MDPs) and propose a method (Greedy Risk-Sensitive Value Iteration, G-RSVI) to compute risk-aware policies for finite horizon MDPs
- Incorporate succinct measures of risk from financial literature to the assessment of algorithmic recourse; these measures are Value at Risk (Holton 2013) and Conditional Value at Risk (Rockafellar and Uryasev 2000).
- Evaluate the policies with different risk profiles computed by G-RSVI on two real-world datasets (UCI Adult Income, German Credit), considering risk measures, sparsity, and proximity; demonstrate that the latter do not implicitly factor in risks.
- Initiate an investigation into gender disparities in terms of risk exposure in the aforementioned datasets.

Motivating Example

To better illustrate the concept of SafeAR with risk-aware policies, consider the following motivating example. 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 scenario, the action costs are in terms of time units (months), each action taken has a probability of success, and failure could transition into a less favorable state. Let us look at three recourse policies that could be given to the applicant illustrated in Figure 1:

- *Policy A: Nearest CE, 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. Therefore, the expected time cost would be much more than 1 month.

- *Policy B: Risk-Neutral, Expected Cost 1.5*. This policy recommends 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 becoming unemployed, from which the cost to recover would be much higher. The expected total cost when considering probabilities is the lowest among the three policies. If optimizing for expected cost alone, this policy would be found and might lead the applicant to a worse situation in which they require 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, Expected Cost 2.2*. This policy provides a safer policy to the applicant, where failures do not lead to a worse situation. The actions for this recourse are *Improve-Education-in-Part-Time* and then *Increase-Savings*. The risk associated with this policy is lower than Policy B, but it has a higher expected cost. Policy C might not be found by methods that minimize proximity, as improving educational background could be considered a larger change than getting a higher-paying job.

With the presence of uncertainties, different recourse trajectories (outcomes) might be encountered when following the recourse policy. Figure 2 visualizes the probabilities of possible outcomes and their associated costs for this example. With the risk-averse Policy C, the applicant is able to receive the desired outcome in 3 time-steps (cost) with probability 98%, and the risk of it taking more than 3 is much less than Policy A or B, even if the expected cost is higher than Policy B. Computing such diverse policies in terms of risk and surfacing the risk information to empower the affected individual is the motivation behind SafeAR.

Related Work

Existing algorithmic recourse methods (Karimi, Schölkopf, and Valera 2021) can be grouped into three categories: one involves finding the nearest CEs as the smallest changes to the individual's feature vector. Solutions focus on *proximity* (Wachter, Mittelstadt, and Russell 2017), *sparsity*, and *diversity* (Karimi et al. 2020; Van Looveren and Klaise 2021; Mothilal, Sharma, and Tan 2020; Kanamori et al. 2020) using multi-objective optimization (Dandl et al. 2020) and decision trees (Kanamori et al. 2022). Also, generative algorithms (Joshi et al. 2019; Barredo-Arrieta and Del Ser 2020) are used to ensure *plausibility*, by generating CEs within data distributions. These methods 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 (Ustun, Spangher, and Liu 2019) or by providing a path over the feature-space along dense regions of the data manifold (Poyiadzi et al. 2020) considering *feasibility* and *actionability*. Methods also incorporate *causality* through structural causal models (SCMs) (Shimizu et al. 2011) to explicitly model inter-variable causal relationships (Mahajan, Tan, and Sharma 2019) and provide an ordered sequence of CEs (Kanamori et al. 2021; Naumann and Ntoutsi 2021). Lastly, robust recourse methods (Upadhyay, Joshi, and Lakkaraju 2021; Nguyen, Bui, and Nguyen 2023) 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 (Singh et al. 2023). "ReLAX" (Chen et al. 2022) generates recourse plans by deep reinforcement learning but under deterministic feature transitions, ignoring uncertainties and thus risk. FASTAR (Verma, Hines, and Dickerson 2022) presents a framework that translates an algorithmic recourse problem into a discounted MDP and demonstrates comparable recourse performance as CE methods. Although FASTAR models uncertainties, it only optimizes for the expected cost and does not incorporate any risk measures. Our first method for SafeAR (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) (Sobel 1994; Prashanth and Ghavamzadeh 2016). There are also other RL methods that factor in risks in policies (Fei, Yang, and Wang 2021; Howard and Matheson 1972; Borkar 2010; Chow et al. 2015b). 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.

Safe Algorithmic Recourse

Algorithmic Recourse Problem Statement

Let $f : \mathcal{X} \to \mathcal{Y}$ be a decision function operationalized by a 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.

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 $\langle S, A, T, R, H \rangle$.

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, "resting heart rate" can be a feature in a health insurance premiums calculator f, but "average calories burned" is not. However, the latter may be in the recourse state, as it is directly affected by actions (e.g., *Exercise*), and in turn can have a causal effect on "resting heart rate". Using only the same features as in f may not be adequate for computing recourse policies, as they may not cover the states and actions that a person actually has to change during the recourse. This gives us a reason to expect separate action, transition, and cost models for recourse, rather than assuming they can be extracted from the data used in f. Such formulation also enables the recourse to provide personalized recommendations to individuals, as advocated for in (Venkatasubramanian and Alfano 2020).

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 (Karimi et al. 2022): 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 a 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. $r(s, a, s'; f)$ tells us the cost incurred to the recipient when the transition (s, a, s') occurs during the recourse. Additionally, when the ML model f gives the favorable outcome in a state ($f(s) = y^+$), then no more actions are needed in the recourse. To capture this, we add a zero-cost action in all favorable (goal) states, transitioning to the same state.

As for the real-world semantics of the cost, it can be a combined measure of multiple factors such as elapsed time, material expenses, opportunity cost, etc. The cost may be averaged across a group or tailored for each person, which requires domain knowledge—so do the feasible actions and the transitions. CE methods such as DiCE (Mothilal, Sharma, and Tan 2020) and FACE (Poyiadzi et al. 2020) also require domain knowledge to design distance (cost) functions, where the function $r(.)$ can be defined in terms of how much the state changes by an action using *sparsity* of feature changes and *proximity* of the recipient's state changes over pre-defined distance functions.

Horizon (H) and Recourse Policies Horizon H is the maximum number of steps in the finite horizon MDP, and $h := [1 : H]$ is the step number over the horizon. 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 $\pi = (\pi_1 \dots \pi_H)$, where $\pi_i := S \rightarrow A$ maps each state to an action for each step i.

SafeAR Methodology

We present a method to compute risk-averse recourse policies and measures to evaluate the risk for SafeAR.

Greedy Risk-Sensitive Value Iteration

We present a greedy algorithm to compute risk-averse policies by incorporating cost variance. 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 recourse cost accrued over the horizon H from a rollout obtained by following a policy π starting at state s and step h. Since outcomes of actions are probabilistic, the total recourse cost is a distribution. 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 also considers the variance in the total recourse cost to manage risk and seeks to find a policy π to maximize the following value function for each state s starting at the first step $h = 1$:

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

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 the 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 end 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} is the value computed in the previous step. The risk-sensitive action value or Q-value $Q_h(s, a)$ at step h is defined as:

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

If only optimizing the expected reward, this procedure would find the optimal policy because the optimal substructure assumption for dynamic programming holds. However, G-RSVI does not guarantee 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

Algorithm 1: G-RSVI

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

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

the state and horizon space. 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 in our algorithm. For example, one can also penalize Lower Partial Standard Deviation (LPSD) (Karagyozova 2016) (Appendix A.8) of lower-than-average values (negative) to avoid higher-than-average costs, as greater negative values indicate higher costs. Equation 3 helps us compare against risk-neutral policies that optimize for expected value only, by setting $\beta = 0$. We leave the analysis of different risk-sensitive algorithms for recourse to future work.

Risk Measures for Recourse Policies

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

Success Rate (ρ_H): It estimates the probability of success within the 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. ρ_H is not equivalent to *validity*, which only determines whether a feature instance 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$_\alpha$): VaR (Holton 2013) is to provide a succinct probabilistic guarantee on the recourse policy cost. We evaluate VaR (Holton 2013) of the recourse cost to answer the question "What is the highest cost at a given level of cumulative probability (confidence level)". For example, 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$_\alpha$ is computed as:

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

Conditional Value at Risk (CVaR$_\alpha$): CVar (Rockafellar and Uryasev 2000) is a complementary measure to VaR and tells us the expected worst-case cost when the cost exceeds the threshold given by VaR$_\alpha$ value. For example, CVaR$_{95} = 8.4$ means that when the cost exceeds the 95-percentile cost, the average cost for those cases is 8.4. It is computed as:

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

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) (Becker and Kohavi 1996) and German Credit Dataset (GCD) (1000 data points) (Hofmann 1994) 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 (Kaggle 2016) with 9 features. To process the datasets for G-RSVI, we convert continuous feature values into discrete values (details included in Appendix A.1). We then train random forest classifiers for both datasets. Dataset features, feature state dimensions, and classifier accuracies are reported in Table 2. G-RSVI does not access the dataset when computing the policy and only uses the trained ML model to indicate whether the desired outcome is received.[2]

Transitions and Rewards 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 (Verma, Hines, and Dickerson 2022), 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 to *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 more details on the model values, we refer the reader to Appendix A.2 for an exhaustive list of model transition probabilities and costs. Results from a different model using the same qualitative assumptions are

[2] All supplemental materials (appendices and code implementations) are available through arxiv.org/abs/2308.12367.

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 example instances sampled from the datasets. n denotes the number of instances used for evaluation. For each risk-aversion level, we compute a recourse policy and apply it to each instance (initially with the undesired outcome) in the dataset. Due to the uncertainties in action outcomes, we estimate the average of each measure using simulated 100 trials. Then, the average across all instances for each metric is computed. The best metric values among the policies are highlighted in bold.

Dataset (#Feat.)	Immutable Features	#States	ML Model (Accuracy)
AID (8)	Gender, Race, Marital Status	57600	Rand.Forest (0.81)
GCD (9)	Gender, Purpose, Credit Amount	147456	Rand.Forest (0.76)

Table 2: Dataset Overview

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, CE methods do not naturally allow probabilities into the formulation, and FASTAR (Verma, Hines, and Dickerson 2022) 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 $\beta = 0.25, 0.50, 0.75$ 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 (L_0 distance) and proximity (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 (sampled 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, 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, 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 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 Risk Disparity Inspired by prior work that investigated recourse fairness (Sharma, Henderson, and Ghosh 2020; Haldar, Cunningham, and Ferhatosmanoglu 2022; von Kügelgen et al. 2022; Raimondi, Lawrence, and Chockler 2022), we now look at the disparity that may exist in risk

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. The male group is exposed to less risk under the policy with the same risk-aversion level.

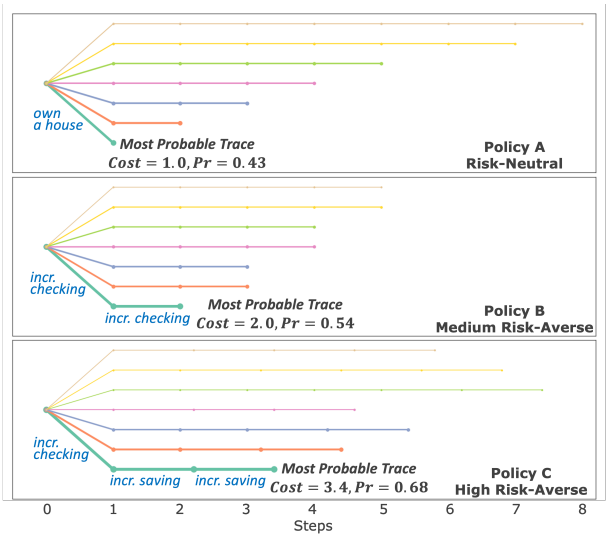

Figure 3: Policy visualization Example in German Credit.

measures across two gender groups (male and female) provided in AID and GCD. Table 3 reports the risk measures (averaged) for female and male groups in both datasets. The p-values for statistical significance of the difference between the groups are provided in Appendix A.4. 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 for both datasets, meaning that for the policy with the same risk-aversion level given to females and males, we expect females would get higher variance in cost (σ^2_{cost}), higher cost at the VaR confidence level of $\alpha = 80, 95$, and higher costs in the expected worst case scenarios (CVaR) for those confidence levels. In AID, 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. The only difference is the decisions made by the model f for different groups, which affects the number of steps to reach the favorable state ($f(x) = y^+$). This is something that recourse providers may want to keep in mind, and it motivates further discussion on risk disparity in algorithmic recourse.

Discussion and Conclusions

The motivation behind our Safe Algorithmic Recourse (SafeAR) is to offer recourse policies with different risk profiles. This enables affected individuals to be aware of the risks and helps them make an informed decision 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, which motivates further studies on recourse fairness in terms of risk exposure.

Ethical Statement

SafeAR would work best with personalized action costs and action success likelihoods for an individual (no personal data was collected or used in this work). If done that way, it would effectively mean asking for personal information. However, the personal data can be deleted after computing the recourse policy, as SafeAR only requires individual action-model data for computation. Also, the black-box ML model does not need information on personal models and preferences on actions and transitions.

Acknowledgements

This paper was prepared for informational purposes by the Artificial Intelligence Research group of JPMorgan Chase & Co. and its affiliates ("JP Morgan"), and is not a product of the Research Department of JP Morgan. JP Morgan makes no representation and warranty whatsoever and disclaims all liability, for the completeness, accuracy or reliability of the information contained herein. This document is not intended as investment research or investment advice, or a recommendation, offer or solicitation for the purchase or sale of any security, financial instrument, financial product or service, or to be used in any way for evaluating the merits of participating in any transaction, and shall not constitute a solicitation under any jurisdiction or to any person, if such solicitation under such jurisdiction or to such person would be unlawful.

References

Barredo-Arrieta, A.; and Del Ser, J. 2020. Plausible Counterfactuals: Auditing Deep Learning Classifiers with Realistic Adversarial Examples. *International Joint Conference on Neural Networks (IJCNN)*, 1–7.

Bäuerle, N.; and Rieder, U. 2014. More risk-sensitive Markov decision processes. *Mathematics of Operations Research*, 39(1): 105–120.

Beam, A. L.; and Kohane, I. S. 2018. Big data and machine learning in health care. *Jama*, 319(13): 1317–1318.

Becker, B.; and Kohavi, R. 1996. Adult. UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C5XW20.

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

Chen, Z.; Silvestri, F.; Wang, J.; Zhu, H.; Ahn, H.; and Tolomei, G. 2022. ReLAX: Reinforcement Learning Agent Explainer for Arbitrary Predictive Models. *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM '22)*, 252–261.

Chow, Y.; Tamar, A.; Mannor, S.; and Pavone, M. 2015a. Risk-sensitive and robust decision-making: a cvar optimization approach. *Advances in neural information processing systems*, 28.

Chow, Y.; Tamar, A.; Mannor, S.; and Pavone, M. 2015b. Risk-sensitive and robust decision-making: a cvar optimization approach. *Advances in neural information processing systems*, 28.

Dandl, S.; Molnar, C.; Binder, M.; and Bischl, B. 2020. Multi-Objective Counterfactual Explanations. *Parallel Problem Solving from Nature – PPSN XVI*, 448–469.

Fei, Y.; Yang, Z.; and Wang, Z. 2021. Risk-Sensitive Reinforcement Learning with Function Approximation: A Debiasing Approach. *Proceedings of the 38th International Conference on Machine Learning*, 139: 3198–3207.

Guidotti, R. 2022. Counterfactual explanations and how to find them: literature review and benchmarking. *Data Mining and Knowledge Discovery*.

Haldar, A.; Cunningham, T.; and Ferhatosmanoglu, H. 2022. RAGUEL: Recourse-Aware Group Unfairness Elimination. *Proceedings of the 31st ACM International Conference on Information and Knowledge Management*, 666–675.

Hofmann, H. 1994. Statlog (German Credit Data). UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C5NC77.

Holton, G. A. 2013. *Value-at-Risk: Theory and Practice, Second Edition*. Online.

Howard, R. A.; and Matheson, J. E. 1972. Risk-sensitive Markov decision processes. *Management science*, 18(7): 356–369.

Joshi, S.; Koyejo, O.; Vijitbenjaronk, W.; Kim, B.; and Ghosh, J. 2019. Towards Realistic Individual Recourse and Actionable Explanations in Black-Box Decision Making Systems. *arXiv*.

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

Kanamori, K.; Takagi, T.; Kobayashi, K.; and Arimura, H. 2020. DACE: Distribution-Aware Counterfactual Explanation by Mixed-Integer Linear Optimization. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, 2855–2862.

Kanamori, K.; Takagi, T.; Kobayashi, K.; and Ike, Y. 2022. Counterfactual Explanation Trees: Transparent and Consistent Actionable Recourse with Decision Trees. *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, 151: 1846–1870.

Kanamori, K.; Takagi, T.; Kobayashi, K.; Ike, Y.; Uemura, K.; and Arimura, H. 2021. Ordered Counterfactual Explanation by Mixed-Integer Linear Optimization. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(13): 11564–11574.

Karagyozova, T. 2016. 4.4 Lower Partial Standard Deviation. https://ecampusontario.pressbooks.pub/econ/chapter/4-4-lower-partial-standard-deviation/. Accessed: 2023-12-10.

Karimi, A.-H.; Barthe, G.; Balle, B.; and Valera, I. 2020. Model-Agnostic Counterfactual Explanations for Consequential Decisions. *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, 108: 895–905.

Karimi, A.-H.; Barthe, G.; Schölkopf, B.; and Valera, I. 2022. A Survey of Algorithmic Recourse: Contrastive Explanations and Consequential Recommendations. *ACM Computing Surveys*, 55(5): 1–29.

Karimi, A.-H.; Schölkopf, B.; and Valera, I. 2021. Algorithmic Recourse: From Counterfactual Explanations to Interventions. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 353–362.

Li, H.; Cao, Y.; Li, S.; Zhao, J.; and Sun, Y. 2020. XGBoost model and its application to personal credit evaluation. *IEEE Intelligent Systems*, 35(3): 52–61.

Mahajan, D.; Tan, C.; and Sharma, A. 2019. Preserving Causal Constraints in Counterfactual Explanations for Machine Learning Classifiers. *arXiv:1912.03277*.

Mothilal, R. K.; Sharma, A.; and Tan, C. 2020. Explaining Machine Learning Classifiers through Diverse Counterfactual Explanations. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 607–617.

Naumann, P.; and Ntoutsi, E. 2021. Consequence-Aware Sequential Counterfactual Generation. *Machine Learning and Knowledge Discovery in Databases. Research Track*, 12976: 682–698.

Nguyen, D.; Bui, N.; and Nguyen, V. A. 2023. Distributionally Robust Recourse Action. *The Eleventh International Conference on Learning Representations*.

Poyiadzi, R.; Sokol, K.; Santos-Rodriguez, R.; De Bie, T.; and Flach, P. 2020. FACE: Feasible and Actionable Counterfactual Explanations. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 344–350.

Prashanth, L.; and Ghavamzadeh, M. 2016. Variance-constrained actor-critic algorithms for discounted and average reward MDPs. *Machine Learning*, 105: 367–417.

Raimondi, F. E. D.; Lawrence, A. R.; and Chockler, H. 2022. Equality of Effort via Algorithmic Recourse. arXiv:2211.11892.

Rockafellar, R. T.; and Uryasev, S. 2000. Optimization of conditional value-at risk. *Journal of Risk*, 2(3): 21–41.

Sharma, S.; Henderson, J.; and Ghosh, J. 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*, 166–172.

Shimizu, S.; Inazumi, T.; Sogawa, Y.; Hyvärinen, A.; Kawahara, Y.; Washio, T.; Hoyer, P. O.; and Bollen, K. 2011. DirectLiNGAM: A Direct Method for Learning a Linear Non-Gaussian Structural Equation Model. *Journal of Machine Learning Research*, 12(33): 1225–1248.

Singh, R.; Miller, T.; Lyons, H.; Sonenberg, L.; Velloso, E.; Vetere, F.; Howe, P.; and Dourish, P. 2023. Directive Explanations for Actionable Explainability in Machine Learning Applications. *ACM Transactions on Interactive Intelligent Systems*, 34.

Sobel, M. J. 1994. Mean-variance tradeoffs in an undiscounted MDP. *Operations Research*, 42(1): 175–183.

ul Hassan, C. A.; Iqbal, J.; Hussain, S.; AlSalman, H.; Mosleh, M. A.; and Sajid Ullah, S. 2021. A computational intelligence approach for predicting medical insurance cost. *Mathematical Problems in Engineering*, 2021: 1–13.

Upadhyay, S.; Joshi, S.; and Lakkaraju, H. 2021. Towards Robust and Reliable Algorithmic Recourse. *Advances in Neural Information Processing Systems*.

Ustun, B.; Spangher, A.; and Liu, Y. 2019. Actionable Recourse in Linear Classification. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 10–19.

Van Looveren, A.; and Klaise, J. 2021. Interpretable Counterfactual Explanations Guided by Prototypes. *Machine Learning and Knowledge Discovery in Databases. Research Track*, 650–665.

Venkatasubramanian, S.; and Alfano, M. 2020. The philosophical basis of algorithmic recourse. *Proceedings of the 2020 conference on fairness, accountability, and transparency*, 284–293.

Verma, S.; Hines, K.; and Dickerson, J. P. 2022. Amortized Generation of Sequential Algorithmic Recourses for Black-Box Models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(8): 8512–8519.

Von Kügelgen, J.; Karimi, A.-H.; Bhatt, U.; Valera, I.; Weller, A.; and Schölkopf, B. 2022. On the fairness of causal algorithmic recourse. *Proceedings of the AAAI conference on artificial intelligence*, 36(9): 9584–9594.

von Kügelgen, J.; Karimi, A.-H.; Bhatt, U.; Valera, I.; Weller, A.; and Schölkopf, B. 2022. On the Fairness of Causal Algorithmic Recourse. *Proceedings of the 36th AAAI Conference on Artificial Intelligence*, 9: 9584–9594.

Wachter, S.; Mittelstadt, B.; and Russell, C. 2017. Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harv. JL & Tech.*, 31: 841.