RISE: Robust Individualized Decision Learning with Sensitive Variables

Xiaoqing Tan [1] Zhengling Qi [2] Christopher W. Seymour [1 3] Lu Tang [1]

Abstract

This paper introduces RISE, a robust individualized decision learning framework with sensitive variables, where sensitive variables are collectible data and important to the intervention decision, but their inclusion in decision making is prohibited due to reasons such as delayed availability or fairness concerns. The convention is to ignore these sensitive variables in learning decision rules, leading to significant uncertainty and bias. To address this, we propose a decision learning framework to incorporate sensitive variables during *offline* training but do not include them in the input of the learned decision rule during model deployment. Specifically, from a causal perspective, the proposed framework intends to improve the worst-case outcomes of individuals caused by sensitive variables that are unavailable at the time of decision. Unlike most existing literature that uses mean-optimal objectives, we propose a robust learning framework via finding a newly defined quantile- or infimum-optimal decision rule. The reliable performance of the proposed method is demonstrated through synthetic experiments and three real-data applications.

1. Introduction

Recently, there has been a widespread interest in developing methodology for individualized decision rules (IDRs) based on observational data. When deriving IDRs, some collectible data are important to the intervention decision, while their inclusion in decision making is prohibited due to reasons such as delayed availability or fairness concerns. For example, sensitive characteristics of subjects regarding their income, sex, race and ethnicity may not be appropriate to be used directly for decision making due to fairness concerns. In the medical field especially for patients in severe life-threatening conditions such as sepsis, timely bedside intervention decisions have to be made before lab measurements are ordered, assayed and returned to the attending physicians. However, due to the delayed availability of lab results, most of the decisions are made with great uncertainty and bias due to partial information at hand. We define *sensitive variables* as variables whose inclusion into decision rules is prohibited. The formal definition of sensitive variables will be given in Section 2.

In this work, we propose RISE (**R**obust **I**ndividualized decision learning with **SE**nsitive variables), a robust IDR framework to improve the outcome of individuals when there are informative yet sensitive variables that are either not available or prohibited from using during IDR deployment. The main question of interest is whether the learned IDR could yield similar outcomes across all realizations of the sensitive variables. To achieve this, we propose to estimate the optimal IDR by optimizing a quantile- or infimum-based objective, respectively, for continuous or discrete sensitive variables. Our idea falls along the lines of work that considers algorithmic fairness (Dwork et al., 2012) while extending it to the setting of causal inference (Rubin, 2005) in the sense that decisions are driven by causality rather than a general utility function. We show in our empirical analyses that this leads to fairer and safer real-life decisions with little sacrifice of the overall performance. This optimization problem is then shown to be equivalent to a weighted classification problem where *most existing statistical and machine learning classifiers can be readily applied*.

Assuming that a large outcome value is preferable, optimal IDRs are traditionally derived through maximizing the mean outcome of the sample population. In this paper, we are interested in a specific yet broadly applicable setting of learning that involves sensitive variables. We consider offline learning where sensitive variables are collected and *can be used in training the IDRs*, but they *cannot be used as input in the resulting IDRs*. This is a setting commonly considered in the fairness and privacy-related literature for classification (e.g., Kamiran & Calders (2012)), but not from a causal standpoint. We defer a detailed discussion of related work to Appendix A. When there exist important variables that are simply left out from training, the estimated

[1]University of Pittsburgh, Pittsburgh, PA, USA [2]George Washington University, Washington, DC, USA [3]University of Pittsburgh Medical Center, Pittsburgh, PA, USA. Correspondence to: Lu Tang <lutang@pitt.edu>.

ICML 2022 Workshop on Responsible Decision Making in Dynamic Environments, Baltimore, Maryland, USA, 2022. Copyright 2022 by the author(s).

IDR will be biased. This bias can be removed if all important variables are used during training, which we will show in Section 2.1 a mean-optimal approach. The optimal action maximizes the mean outcome where the mean is taken over the sensitive variables, conditioning on other variables. This method, however, has no control of the disparity in sensitive variables. Subjects with different sensitive values may report large outcome differences, hence unfairly or unsafely treated. Therefore, objective functions with robustness guarantees for sensitive variables are preferred, since they offer protection to subjects in the lower tail of the outcome distribution with regards to the sensitive values.

For illustration, we consider a toy example with binary actions, $A \in \{-1, 1\}$. We note that the decision can only be made based on the variable X whereas S is a sensitive variable. The setup is shown in Table 1 and partial simulation results are shown in Table 2. The details can be found in Section 3.1 under Example 1. We consider *vulnerable subjects* as those with low outcome values, as highlighted in red in Table 1 (A full definition will be given in Section 2.2). For $X \leq 0.5$, the traditional mean-optimal rule assigns action $A = 1$ as it achieves the largest average reward across $S = 0$ and $S = 1$. However, this action results in great harm for subjects with $S = 1$ as they could get the worst expected outcome of 0. On the contrary, the proposed RISE improves the worst-case outcome by assigning $A = -1$, protecting the vulnerable subjects. Likewise, for $X > 0.5$, the mean-optimal rule assigns $A = -1$ while the proposed rule assigns $A = 1$ protecting those with $S = 0$ that could have experienced an outcome of 5. Compared to the mean-optimal rule, the proposed RISE achieves a much larger reward among vulnerable subjects while maintaining a comparable overall expected reward. The worst-case outcomes of the rule by the proposed RISE are colored in blue.

Table 1: Toy example setup.

	$X \leq 0.5$		$X > 0.5$	
$E(Y\|X, S, A)$	$S = 0$	$S = 1$	$S = 0$	$S = 1$
$A = -1$	11	13	5	27
$A = 1$	30	0	15	13

Table 2: Toy example results.

	Average reward	
	Overall	Vulnerable
Mean-optimal rule	**14.4**	7.1
RISE	13	**14**

Main contributions. *Methodology-wise, 1)* we propose a novel framework, RISE, to handle sensitive variables in causality-driven decision making. Robustness is introduced to improve the worst-case outcome caused by sensitive variables, and as a result, it reduces the outcome variation across subjects. The latter is directly associated with fairness and safety in decision making. To the best of our knowledge, we are among the first to propose a robust-type fairness criterion under causal inference. *2)* We introduce a classification-based optimization framework that can easily leverage most existing classification tools catered to different functional classes, including state-of-the-art random forest, boosting, or neural network models. *Application-wise, 3)* the consideration of sensitive variables in decision learning is important to applications in policy, education, healthcare, etc. Specifically, we illustrate the application of RISE using three real-world examples from fairness and safety perspectives where robust decision rules are needed, across which we have observed robust performance of the proposed approach. From a fairness perspective, we consider a job training program where age is considered as a sensitive variable. From a safety perspective, we consider two applications to healthcare where lab measurements are considered as sensitive variables.

2. Robust Decision Learning Framework with Sensitive Variables

2.1. Preliminaries

Notation. We let random variables be represented by upper-case letters, and their realizations be represented by lower-case letters. Suppose there are n independent subjects sampled from a given population. For subject i, let $A_i \in \{-1, 1\}$ denote a binary treatment assignment and Y_i denote the corresponding outcome. Without loss of generality, we assume a larger value of outcome is desirable. Under the potential outcomes framework (Rubin, 1978; Splawa-Neyman et al., 1990), let $Y_i(-1)$ be the potential outcome had the subject been assigned to control and $Y_i(1)$ be the potential outcome had the subject been assigned to treatment. Let $X_i \in \mathbb{X}$ be the feature vector and, for now, S_i be a single sensitive variable. Extension to multiple sensitive variables is presented in Section 2.4. We consider $S \in \mathbb{S}$ where $\mathbb{S} = \{1, \ldots, K\}$ if S is discrete and $\mathbb{S} = \mathbb{R}$ if S is continuous.

Definition of sensitive variables. We define sensitive variables that are important to the intervention decision, but their inclusion in decision making is prohibited. Formally, consider variables X and S that are both available during model training and are both determinants of conditional average treatment effect (Rubin, 1974). While X and S may be both involved in training, the derived decision rule $d(\cdot)$ precludes the input of S due to sensitive concerns. Hence, the derived IDR is only a function with the form $d(X) : \mathbb{X} \rightarrow \mathbb{A}$. Following the above definition, we consider an offline learning framework where sensitive variables are collected and can be used in obtaining the IDRs, but they cannot be used in the resulting IDRs. We defer the required causal assumptions and related discussion to Appendix B.

Naive approaches that omit sensitive variables. When S is not available for future deployment, a naive approach is to maximize $E_X\{E(Y|X, A = d(X))\}$ over d using (X, A, Y) during training procedure. This approach introduces bias in the estimation of potential outcomes and leads to a suboptimal IDR due to the unmeasured confounder S.

Mean-optimal approaches that use sensitive variables. It is thus important that one includes S into the training procedure. For example, if we consider the value function framework (i.e., expected outcome) used by most existing works such as Manski (2004); Qian & Murphy (2011), we can show that

$$\begin{aligned} E\{Y(d)\} &= E_{X,S}\big[E(Y(d)|X, S)\big] \\ &= E_X\big[E_{S|X}\{E(Y(d)|X, S)\}\big] \\ &= E_X\big[E_{S|X}\{E(Y|X, S, A = d(X))\}\big] \quad (1)\\ &\neq E_X\big[E(Y|X, A = d(X))\big], \end{aligned}$$

where the third equality in (1) holds by Assumption 1 and the last inequality also indicates the naive approaches without using S will in general fail. Then one valid approach is to maximize $E_X\big[E_{S|X}\{E(Y|X, S, A = d(X))\}\big]$ over d using (X, S, A, Y). The optimal IDR under this criterion is, for every $X \in \mathbb{X}$, $\tilde{d}(X) \in \text{sign}(E_{S|X}\{E(Y|X, S, A = 1)\} - E_{S|X}\{E(Y|X, S, A = -1)\})$, which guarantees to find the treatment that maximizes the conditional expected outcome given each X by averaging out the effect of the sensitive variable S. The mean-optimal approaches, however, fail to control the disparities across realizations of the sensitive variables due to the integration over S, which may lead to unsatisfactory decisions to certain subgroups, as illustrated in the toy example in Section 1.

2.2. Robust Optimality with Sensitive Variables

Driven by the limitation of existing approaches, our goal is to derive a robust decision rule that maximizes the worst-case scenarios of subjects when some sensitive information is not available at the time of deploying the decision rule. Specifically, our robust decision learning framework draws decisions based on individuals' available characteristics summarized in the vector X without the sensitive variable S, while improving the worst-case outcome of subjects in terms of the sensitive variable in the population. Formally, given a collection $\mathbb{D}$ of all treatment decision rules depending only on X, the proposed RISE approach estimates the following IDR, which is defined as

$$d^* \in \arg\max_{d\in\mathbb{D}} E_X\big[G_{S|X}\{E(Y|X, S, A = d(X))\}\big], \quad (2)$$

where $G_{S|X}(\cdot)$ could be chosen as some risk measure for evaluating $E(Y|X, S, A = d(X))$ for each $S \in \mathbb{S}$. Examples include variance, conditional value at risk, quantiles, etc. In this paper, we consider $G_{S|X}$ as the conditional quantiles (for a continuous S) or the infimum (for a discrete S) over $\mathbb{S}$.

Specifically, for a discrete S, $G_{S|X}$ is consider as an infimum operator of $E(Y|X, S, A = d(X))$ over S. We thus aim to find $d^* \in \text{argmax}_{\mathbb{D}} E_X\big[\inf_{s\in\mathbb{S}}\{E(Y|X, S = s, A = d(X))\}\big]$, where $\inf$ is the infimum taken with respect to $E(Y|X, s, A = d(X))$ over $s \in \mathbb{S}$. This implies that for a given X, $d^*(X)$ assigns the treatment that yields the best worst-case scenario among all possible values of S for every $X \in \mathbb{X}$, or equivalently, $d^*(X) \in \text{sign}(\inf_{s\in\mathbb{S}}\{E(Y|X, S = s, A = 1) - \inf_{s\in\mathbb{S}}\{E(Y|X, S = s, A = -1)\})$. For a continuous S, we consider $G_{S|X}\{E(Y|X, S, A = d(X))\}$ as $Q^{\tau}_{S|X}\{E(Y|X, S, A = d(X))\}$, which is the τ-th quantile of $\{E(Y|X, S, A = d(X))\}$ and $\tau \in (0, 1)$ is the quantile level of interest. Specifically, $Q^{\tau}_{S|X}\{E(Y|X, S, A = d(X))\} = \inf\{t : F(t) \geq \tau\}$ with F denoting the conditional distribution function of $E(Y|X, S, A = d(X))$ given X and d. Note the randomness behind $E(Y|X, S, A = d(X))$ given X and d is fully determined by the sensitive variable S. Then optimal IDR under this criterion is defined as $d^* \in \text{argmax}_{\mathbb{D}} E_X\big[Q^{\tau}_{S|X}\{E(Y|X, S, A = d(X))\}\big]$. This implies that for a given X, $d^*(X)$ assigns a treatment that yields the largest τ-th quantile of the outcome over the distribution related to S, or equivalently, $d^*(X) \in \text{sign}(\{Q^{\tau}_{S|X}\{E(Y|X, S, A = 1)\} - Q^{\tau}_{S|X}\{E(Y|X, S, A = -1)\})$. We let $\tau = 0.25$ throughout and suppress τ for simplicity. Results on varying the value of τ is provided in Appendix; see Section 3.1 for details.

Identifying vulnerable subjects. Our RISE framework provides a natural way to define *vulnerable groups*. Specifically, for a discrete S, if $\inf_S\{E(Y|X, S, A = 1)\} > \inf_S\{E(Y|X, S, A = 0)\}$, then $\arg\inf_S\{E(Y|X, S, A = 0)\}$ is vulnerable given X, otherwise $\arg\inf_S\{E(Y|X, S, A = 1)\}$ is vulnerable. In other words, the vulnerable subjects are those in the worst-off group that needs protection. Similarly, for a continuous S, if $Q_S\{E(Y|X, S, A = 1)\} > Q_S\{E(Y|X, S, A = 0)\}$, then the set $\{S : E(Y|X, S, A = 0) \leq Q_S\{E(Y|X, S, A = 0)\}\}$ defines the vulnerable subjects given X, otherwise this group is defined as $\{S : E(Y|X, S, A = 1) \leq Q_S\{E(Y|X, S, A = 1)\}\}$.

2.3. Estimation and Algorithm

Here we provide a transformation of the proposed RISE from an optimization problem to a weighted classification problem. There are several advantages to this conversion: 1) The optimization problem defined in (2) involves a nonsmooth and nonconvex objective function that could lead to computational challenges. 2) With multiple powerful statistical and machine learning toolbox to choose from, a classification problem can be more readily solved in practice. Hyperparameter tuning and model selection could be conducted to further boost performance. 3) Compared to a direct optimization of (2), a classification-based optimizer allows the use of off-the-shelf software packages that can be tailored to different functional classes or incorporate different properties such as model sparsity.

With Proposition 1 and Proposition 2, we have transformed the optimization problem (2) into a weighted classification problem (4) where for subject i with features x_i, the true label is $\text{sgn}\{g_1(x_i) - g_2(x_i)\}$ and the sample weight is $|g_1(x_i) - g_2(x_i)|$. The estimated optimal decision rule by (4) is then given by $\hat{d}(x) = \text{sgn}\{\hat{f}(x)\}$, where $f(x)$ is a smooth function. Proposition 1 and Proposition 2 along with their proofs are presented in Appendix C. The detailed description of the algorithm as well as modeling and hyperparameter tuning via cross-validations can be found in Appendix D.

2.4. Extension to multiple discrete sensitive variables

For multiple discrete sensitive variables, similar estimation procedure can be conducted as outlined in Section 2.3. Suppose there are L discrete sensitive variables, i.e., $\mathcal{S} = \{S_1, S_2, \ldots, S_L\}$. The inner expectation $E(Y|X, S_1, \ldots, S_L, A)$ can be obtained with a twin model of Y on X and all $\mathcal{S}$ for each treatment level. The infimum over $\mathbb{S}$ is obtained by finding the minimum iterating space of possible parameter values for each sensitive variable. See Section 3.2 for an example of using multiple discrete sensitive variables.

3. Numerical Studies

In this section, we perform extensive numerical experiments to investigate the merit of robustness of the proposed framework via simulations and three real-data applications. For comparison, we consider the naive and mean-optimal approaches described in Section 2.1. The details can be found in Appendix E. The numerical results demonstrate that the proposed rules achieve a robust objective with sensitive variables unavailable at the time of decision while maintaining comparable mean outcomes.

Evaluation metrics. *1) Objective:* the quantile objective is estimated and reported for a continuous S and the infimum objective is for a discrete S. The objective, when $\tau < 0.5$, (here $\tau = 0.25$) represents the value of the "low performers" among all possible value of S under a given d. *2) Value:* the value function used by the most existing methods. It represents the "average performers". We report the metrics among all subjects and among the potential vulnerable subgroup, respectively.

3.1. Simulation Studies

A detailed description of simulation setups and results can be found in Appendix F. The proposed RISE achieves the largest objectives and improves the value among vulnerable subjects, while maintaining comparative overall values. As demonstrated in the toy example introduced in Section 1, we expect that RISE helps improve the value among the vulnerable subjects while maintaining a comparable overall value. As for the objective, intuitively, the proposed rule is expected to have a larger objective. We also consider for a continuous S different quantile criteria $\tau = 0.1$ and 0.5, respectively, to test the robustness of the proposed RISE. Results show that when τ is small, there is more strength in the proposed method, as the algorithm aims to improve the worst-outcome scenarios. The proposed RISE has the largest gain in objective and value among vulnerable subjects when τ is 0.1, and has similar performance as the compared approaches when τ is 0.5. Besides, we consider a scenario where S is not involved in the data generation of Y, i.e., Assumption 1c is simplified as $\{Y(-1), Y(1)\} \perp A|X$. The objective and value function are similar across all compared approaches, which indicates the robustness of RISE.

3.2. Real-data Applications

We present three real-data examples to showcase the robust performance of RISE. These applications consider either fairness or safety in the context of policy (LaLonde, 1986) and healthcare (Hammer et al., 1996; Seymour et al., 2016) where sensitive variables exist. A detailed description of datasets and results can be found in Appendix G. As expected, RISE is shown to have the largest objective as well as value among vulnerable subjects. The patterns are similar to that in the synthetic experiments in Section 3.1. We provide visualizations by Shapley additive explanations (SHAP) (Lundberg & Lee, 2017) for RISE and the mean-optimal rule, respectively, in Appendix G about feature importance in the final classification models to help interpret important covariates in making the decisions. Overall, the direction of correlations is similar for RISE and the mean-optimal rule, but their ranking of feature importance may be different.

4. Discussion

We have proposed RISE, a robust decision learning framework with a novel quantile-, or infimum-optimal treatment objective intended to improve the worst-case scenarios of individuals when decisions with uncertainty are needed to be made with sensitive yet important information unavailable. Our approach can be applied to a board range of applications, including but not limited to policy, education, healthcare, etc.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. {TensorFlow}: A system for {Large-Scale} machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pp. 265–283, 2016.

Andersen, L. W., Mackenhauer, J., Roberts, J. C., Berg, K. M., Cocchi, M. N., and Donnino, M. W. Etiology and therapeutic approach to elevated lactate levels. *Mayo Clinic Proceedings*, 88(10):1127–1140, 2013. ISSN 0025-6196. doi: https://doi.org/10.1016/j.mayocp.2013.06.012. URL https://www.sciencedirect.com/science/article/pii/S0025619613005557.

Barocas, S., Hardt, M., and Narayanan, A. Fairness in machine learning. *Nips tutorial*, 1:2, 2017.

Beutel, A., Chen, J., Zhao, Z., and Chi, E. H. Data decisions and theoretical implications when adversarially learning fair representations. *arXiv preprint arXiv:1707.00075*, 2017.

Bietti, A., Agarwal, A., and Langford, J. A contextual bandit bake-off. *Journal of Machine Learning Research*, 22(133):1–49, 2021.

Boppana, S. and Goepfert, P. Understanding the CD8 T-cell response in natural HIV control. *F1000Research*, 7, 2018.

Cai, T., Tian, L., Wong, P. H., and Wei, L. Analysis of randomized comparative clinical trial data for personalized treatment selections. *Biostatistics*, 12(2):270–282, 2011.

Chakraborty, B. and Moodie, E. E. *Statistical Methods for Dynamic Treatment Regimes*. Springer, 2013.

Chakraborty, B., Murphy, S., and Strecher, V. Inference for non-regular parameters in optimal dynamic treatment regimes. *Statistical Methods in Medical Research*, 19(3): 317–343, 2010.

Chen, H., Lu, W., Song, R., and Ghosh, P. On learning and testing of counterfactual fairness through data preprocessing. *arXiv preprint arXiv:2202.12440*, 2022.

Chouldechova, A. and Roth, A. A snapshot of the frontiers of fairness in machine learning. *Communications of the ACM*, 63(5):82–89, 2020.

Creager, E., Madras, D., Jacobsen, J.-H., Weis, M., Swersky, K., Pitassi, T., and Zemel, R. Flexibly fair representation learning by disentanglement. In *International Conference on Machine Learning*, pp. 1436–1445. PMLR, 2019.

Dwork, C., Hardt, M., Pitassi, T., Reingold, O., and Zemel, R. Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pp. 214–226, 2012.

Fang, E. X., Wang, Z., and Wang, L. Fairness-oriented learning for optimal individualized treatment rules. *Journal of the American Statistical Association*, 0(0):1–14, 2022. doi: 10.1080/01621459.2021.2008402. URL https://doi.org/10.1080/01621459.2021.2008402.

Feldman, M., Friedler, S. A., Moeller, J., Scheidegger, C., and Venkatasubramanian, S. Certifying and removing disparate impact. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 259–268, 2015.

Ferreira, F. L., Bota, D. P., Bross, A., Mélot, C., and Vincent, J.-L. Serial evaluation of the sofa score to predict outcome in critically ill patients. *The Journal of the American Medical Association*, 286(14):1754–1758, 2001.

Goldberg, Y. and Kosorok, M. R. Q-learning with censored data. *Annals of Statistics*, 40(1):529, 2012.

Hammer, S. M., Katzenstein, D. A., Hughes, M. D., Gundacker, H., Schooley, R. T., Haubrich, R. H., Henry, W. K., Lederman, M. M., Phair, J. P., Niu, M., et al. A trial comparing nucleoside monotherapy with combination therapy in hiv-infected adults with cd4 cell counts from 200 to 500 per cubic millimeter. *New England Journal of Medicine*, 335(15):1081–1090, 1996.

Hashimoto, T., Srivastava, M., Namkoong, H., and Liang, P. Fairness without demographics in repeated loss minimization. In *International Conference on Machine Learning*, pp. 1929–1938. PMLR, 2018.

Henderson, R., Ansell, P., and Alshibani, D. Regret-regression for optimal dynamic treatment regimes. *Biometrics*, 66(4):1192–1201, 2010.

Holland, P. W. Statistics and causal inference. *Journal of the American statistical Association*, 81(396):945–960, 1986.

Howell, M. D., Donnino, M., Clardy, P., Talmor, D., and Shapiro, N. I. Occult hypoperfusion and mortality in patients with suspected infection. *Intensive Care Medicine*, 33(11):1892–1899, 2007.

Huang, X., Choi, S., Wang, L., and Thall, P. F. Optimization of multi-stage dynamic treatment regimes utilizing accumulated data. *Statistics in Medicine*, 34(26):3424–3443, 2015.

Imai, K. and Ratkovic, M. Estimating treatment effect heterogeneity in randomized program evaluation. *The Annals of Applied Statistics*, 7(1):443–470, 2013.

Imbens, G. W. and Angrist, J. D. Identification and estimation of local average treatment effects. *Econometrica*, 62 (2):467–475, 1994. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/2951620.

Joseph, M., Kearns, M., Morgenstern, J. H., and Roth, A. Fairness in learning: Classic and contextual bandits. *Advances in Neural Information Processing Systems*, 29, 2016.

Kamiran, F. and Calders, T. Data preprocessing techniques for classification without discrimination. *Knowledge and Information Systems*, 33(1):1–33, 2012.

Kosorok, M. R. and Moodie, E. E. *Adaptive treatment strategies in practice: planning trials and analyzing data for personalized medicine*. SIAM, 2015.

Krishna, U., Joshi, S. P., and Modh, M. An evaluation of serial blood lactate measurement as an early predictor of shock and its outcome in patients of trauma or sepsis. *Indian Journal of Critical Care Medicine: peer-reviewed, official publication of Indian Society of Critical Care Medicine*, 13(2):66, 2009.

Laber, E. B., Lizotte, D. J., Qian, M., Pelham, W. E., and Murphy, S. A. Dynamic treatment regimes: Technical challenges and applications. *Electronic Journal of Statistics*, 8(1):1225–1272, 2014.

Lahoti, P., Beutel, A., Chen, J., Lee, K., Prost, F., Thain, N., Wang, X., and Chi, E. Fairness without demographics through adversarially reweighted learning. *Advances in Neural Information Processing Systems*, 33:728–740, 2020.

LaLonde, R. J. Evaluating the econometric evaluations of training programs with experimental data. *The American Economic Review*, pp. 604–620, 1986.

Lundberg, S. M. and Lee, S.-I. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 2017.

Manski, C. F. Statistical treatment rules for heterogeneous populations. *Econometrica*, 72(4):1221–1246, 2004.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., and Galstyan, A. A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6):1–35, 2021.

Moodie, E. E., Richardson, T. S., and Stephens, D. A. Demystifying optimal dynamic treatment regimes. *Biometrics*, 63(2):447–455, 2007.

Moodie, E. E., Platt, R. W., and Kramer, M. S. Estimating response-maximized decision rules with applications to breastfeeding. *Journal of the American Statistical Association*, 104(485):155–165, 2009.

Murphy, S. A. Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):331–355, 2003.

Murphy, S. A. An experimental design for the development of adaptive treatment strategies. *Statistics in Medicine*, 24(10):1455–1481, 2005.

Nabi, R. and Shpitser, I. Fair inference on outcomes. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Nabi, R., Malinsky, D., and Shpitser, I. Learning optimal fair policies. In *International Conference on Machine Learning*, pp. 4674–4682. PMLR, 2019.

Neyman, J. Sur les applications de la thar des probabilities aux experiences agaricales: Essay des principle. Excerpts reprinted (1990) in English. *Statistical Science*, 5:463–472, 1923.

Orellana, L., Rotnitzky, A., and Robins, J. M. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, part i: main content. *The International Journal of Biostatistics*, 6(2), 2010a.

Orellana, L., Rotnitzky, A., and Robins, J. M. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, part ii: proofs of results. *The International Journal of Biostatistics*, 6(2), 2010b.

Patil, V., Ghalme, G., Nair, V., and Narahari, Y. Achieving fairness in the stochastic multi-armed bandit problem. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 5379–5386, 2020.

Pearl, J. *Causality*. Cambridge University Press, 2009.

Pessach, D. and Shmueli, E. A review on fairness in machine learning. *ACM Computing Surveys (CSUR)*, 55(3):1–44, 2022.

Qi, Z., Cui, Y., Liu, Y., and Pang, J.-S. Estimation of individualized decision rules based on an optimized covariate-dependent equivalent of random outcomes. *SIAM Journal on Optimization*, 29(3):2337–2362, 2019.

Qi, Z., Pang, J.-S., and Liu, Y. On robustness of individualized decision rules. *Journal of the American Statistical Association*, pp. 1–15, 2022.

Qian, M. and Murphy, S. A. Performance guarantees for individualized treatment rules. *Annals of Statistics*, 39(2): 1180, 2011.

Robins, J., Orellana, L., and Rotnitzky, A. Estimation and extrapolation of optimal treatment and testing strategies. *Statistics in Medicine*, 27(23):4678–4721, 2008.

Robins, J. M. Optimal structural nested models for optimal sequential decisions. In *Proceedings of the Second Seattle Symposium in Biostatistics*, pp. 189–326. Springer, 2004.

Robins, J. M., Hernan, M. A., and Brumback, B. Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5):550–560, 2000.

Rubin, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688, 1974.

Rubin, D. B. Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*, pp. 34–58, 1978.

Rubin, D. B. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322–331, 2005.

Sakr, Y., Jaschinski, U., Wittebole, X., Szakmany, T., Lipman, J., Ñamendys-Silva, S. A., Martin-Loeches, I., Leone, M., Lupu, M.-N., Vincent, J.-L., and ICON Investigators. Sepsis in intensive care unit patients: worldwide data from the intensive care over nations audit. *Open Forum Infectious Diseases*, 5(12):ofy313, 2018.

Sattigeri, P., Hoffman, S. C., Chenthamarakshan, V., and Varshney, K. R. Fairness gan: Generating datasets with fairness properties using a generative adversarial network. *IBM Journal of Research and Development*, 63(4/5):3–1, 2019.

Seymour, C. W., Liu, V. X., Iwashyna, T. J., Brunkhorst, F. M., Rea, T. D., Scherag, A., Rubenfeld, G., Kahn, J. M., Shankar-Hari, M., Singer, M., et al. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (sepsis-3). *The Journal of the American Medical Association*, 315(8):762–774, 2016.

Shankar-Hari, M., Phillips, G. S., Levy, M. L., Seymour, C. W., Liu, V. X., Deutschman, C. S., Angus, D. C., Rubenfeld, G. D., Singer, M., et al. Developing a new definition and assessing new clinical criteria for septic shock: for the third international consensus definitions for sepsis and septic shock (sepsis-3). *The Journal of the American Medical Association*, 315(8):775–787, 2016.

Song, R., Wang, W., Zeng, D., and Kosorok, M. R. Penalized q-learning for dynamic treatment regimens. *Statistica Sinica*, 25(3):901, 2015.

Splawa-Neyman, J., Dabrowska, D. M., and Speed, T. On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, pp. 465–472, 1990.

Tao, Y. and Wang, L. Adaptive contrast weighted learning for multi-stage multi-treatment decision-making. *Biometrics*, 73(1):145–155, 2017.

Thall, P. F., Sung, H.-G., and Estey, E. H. Selecting therapeutic strategies based on efficacy and death in multicourse clinical trials. *Journal of the American Statistical Association*, 97(457):29–39, 2002.

Varshney, K. R. Engineering safety in machine learning. In *2016 Information Theory and Applications Workshop (ITA)*, pp. 1–5. IEEE, 2016.

Vincent, J.-L., Moreno, R., Takala, J., Willatts, S., De Mendonça, A., Bruining, H., Reinhart, C., Suter, P., and Thijs, L. G. The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure. On behalf of the working group on sepsis-related problems of the european society of intensive care medicine. *Intensive Care Medicine*, 22(7):707–710, 1996.

Wager, S. and Athey, S. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018.

Wang, L., Zhou, Y., Song, R., and Sherwood, B. Quantile-optimal treatment regimes. *Journal of the American Statistical Association*, 113(523):1243–1254, 2018a.

Wang, Y., Fu, H., and Zeng, D. Learning optimal personalized treatment rules in consideration of benefit and risk: with an application to treating type 2 diabetes patients with insulin therapies. *Journal of the American Statistical Association*, 113(521):1–13, 2018b.

Watkins, C. J. and Dayan, P. Q-learning. *Machine Learning*, 8(3):279–292, 1992.

Zafar, M. B., Valera, I., Rogriguez, M. G., and Gummadi, K. P. Fairness constraints: Mechanisms for fair classification. In *Artificial Intelligence and Statistics*, pp. 962–970. PMLR, 2017.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. A robust method for estimating optimal treatment regimes. *Biometrics*, 68(4):1010–1018, 2012.

Zhang, B. H., Lemoine, B., and Mitchell, M. Mitigating unwanted biases with adversarial learning. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 335–340, 2018.

Zhang, J. and Bareinboim, E. Fairness in decision-making—the causal explanation formula. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(499):1106–1118, 2012.

Zhao, Y.-Q., Zeng, D., Laber, E. B., Song, R., Yuan, M., and Kosorok, M. R. Doubly robust learning for estimating individualized treatment with censored data. *Biometrika*, 102(1):151–168, 2015.

A. Closely Related Work

Our work focuses on individualized decision rules, which aim at assigning treatment decision based on subject characteristics. Typical model-based methods for deriving IDRs include Q-learning such as Watkins & Dayan (1992); Murphy (2003); Moodie et al. (2007); Chakraborty et al. (2010); Goldberg & Kosorok (2012); Song et al. (2015) and A-learning such as Robins et al. (2000); Murphy (2005) where a model of responses is imposed and the optimal decision rule is obtained by optimizing value function derived from the model. On the other hand, model-free methods such as Robins et al. (2008); Orellana et al. (2010a;b); Zhang et al. (2012); Zhao et al. (2012; 2015) assign values to actions simply through trial and error without pre-specifying a model. Besides, contextual bandit methods (see Bietti et al. (2021) and references therein) test out different actions and automatically learn which one has the most rewarding outcome for a given situation. Other methods include Robins (2004); Moodie et al. (2009); Cai et al. (2011); Henderson et al. (2010); Thall et al. (2002); Imai & Ratkovic (2013); Huang et al. (2015); Tao & Wang (2017). See Chakraborty et al. (2010); Chakraborty & Moodie (2013); Laber et al. (2014); Kosorok & Moodie (2015) and references therein for a comprehensive review. Beyond the field of causal inference, fairness and safety, and robustness are two areas of research that extend well beyond the learning of IDRs. In the following, we provide a review on both, with focus given to work related to causal inference and IDRs.

Fairness and safety in IDRs. The consideration of fairness and safety in machine learning has seen an explosion of interest in the past few years. We refer to Dwork et al. (2012); Varshney (2016); Barocas et al. (2017); Nabi & Shpitser (2018); Hashimoto et al. (2018); Chouldechova & Roth (2020); Mehrabi et al. (2021); Pessach & Shmueli (2022) and references therein for a review of this topic in classification and regression problems. In these work, sensitive variables are also referred to as sensitive or protected attributes. We extend the definition of sensitive variables to include delayed information that is not available at deployment as it is also suitable for this framework.

Among earlier work, preprocessing approaches (Kamiran & Calders, 2012; Feldman et al., 2015; Creager et al., 2019; Sattigeri et al., 2019) and inprocess training approaches (Beutel et al., 2017; Hashimoto et al., 2018; Lahoti et al., 2020) consider disentangling the input X from a known or unknown sensitive variable S so that the transformed X does not contain any information that can be used to trace back to S. Due to the causal nature of IDRs, effect of IDRs cannot be estimated consistently when an informative S is left out and the resulting rule is sub-optimal. This follows from the classic argument that any unmeasured confounding (i.e., S), if not accounted for, would lead to bias. Similar issues persist in contextual bandits (Joseph et al., 2016; Patil et al., 2020). Inside the causal framework, Zhang & Bareinboim (2018); Nabi et al. (2019) extend fairness from prediction to policy learning using causal graphical models by incorporating fairness constraints. Chen et al. (2022) considers counterfactual fairness that seeks to achieve conditional independence of the decisions via data preprocessing. Despite earlier efforts in bringing fairness into the causal framework, most of these approaches only ensure mean zero disparity in S but do not have robustness guarantees in the sense that the variance of the disparity in S is not controlled. Besides, most examples consider a single categorical sensitive variable, but not multiple or continuous ones.

Robustness in IDRs. Recently the statistical literature has witnessed a growing interest in developing robust methods for estimating IDRs. They introduce robustness into the objective function by using quantile-optimal treatment regimes or mean-optimal treatment regimes under certain constraints to improve the gain of individuals at the lower tail of the reward spectrum (Wang et al., 2018a;b; Qi et al., 2022; 2019; Fang et al., 2022). Robustness, in their sense, pertains to the outcome distribution subject to sampling error. When sensitive variables are present, we consider instead the robustness of the outcome distribution subject to the uncertainty due to sensitive variables, providing a more targeted way of ensuring robustness, which is directly related to fairness and safety. Compared to algorithms based on explicit fairness constraints (for example Zafar et al. (2017); Zhang et al. (2018) in classification and Zhang & Bareinboim (2018); Chen et al. (2022) in causal inference) that seek to remove the disparity across different values of S, our method reduces the variance of disparity across S. In addition, constraint-based approaches typically require specialized optimization procedures whereas our approach presents an elegant and systematic way for optimization. To our knowledge, we are the first few to consider decision fairness via a robust objective under the causal framework.

B. Causal Assumptions

Assumption 1. *Assume the following conditions hold:*

(1a) Consistency: $Y = Y(-1)\mathbb{1}(A = -1) + Y(1)\mathbb{1}(A = 1)$.
(1b) Positivity: $0 < Pr(A = 1|X, S) < 1$.

(1c) Unconfoundedness: $\{Y(-1), Y(1)\} \perp A|\{X, S\}$ *and* $\{Y(-1), Y(1)\} \not\perp A|X$.

Assumption 1a is the standard consistency assumption in causal inference and Assumption 1b states that every subject has a nonzero probability of getting the treatment. Assumption 1c states that given X and S, the potential outcomes are independent of the treatment assignments. Besides, the unconfoundedness does not hold when only X is conditional, signifying the importance of S. Under causal settings, Assumption 1c ensures that treatment effects cannot be non-parametrically identifiable without S (See Neyman (1923); Rubin (1974); Holland (1986); Imbens & Angrist (1994); Pearl (2009) and references therein). Approaches such as disentanglement of X from S under supervised learning settings mentioned in Appendix A will introduce bias towards estimating the IDR.

C. Propositions and Proofs

Proposition 1. *Maximizing the objective function in* (2) *is equivalent to maximizing*

$$E_X\big\{\mathbb{1}(d(X)=1)[G_{S|X}\{E(Y|X,S,A=1)\} - G_{S|X}\{E(Y|X,S,A=-1)\}]\big\}.$$

With Proposition 1 and a proper estimator of the outcome model $E(Y|X,S,A)$ using our training data $\mathcal{D}_n = \{X_i, S_i, A_i, Y_i\}_{i=1}^n$, we replace the expectation of Y_i by its estimate $\hat{Y}_i$ and solve the following problem.

$$\arg\max_{d\in\mathbb{D}} n^{-1}\textstyle\sum_{i=1}^n[\mathbb{1}(d(x_i)=1)\{g_1(x_i)-g_2(x_i)\}], \tag{3}$$

where $g_1(x_i) = G_{s|x}\{\hat{Y}_i(x_i, s, a_i = 1)\}$ and $g_2(x_i) = G_{s|x}\{\hat{Y}_i(x_i, s, a_i = -1)\}$. Note that $g_1(x_i) - g_2(x_i)$ may not be positive, which makes Problem (3) difficult to solve. We have the following key proposition 2 to address this issue and transform it into a classification problem.

Proposition 2. *Let* $f(x)$ *to be a smooth function. Maximizing the empirical objective in* (3) *is equivalent to a weighted classification of minimizing*

$$n^{-1}\textstyle\sum_{i=1}^n \mathbb{1}[\operatorname{sgn}\{g_1(x_i)-g_2(x_i)\}\cdot f(x_i)<0]\cdot|g_1(x_i)-g_2(x_i)|, \tag{4}$$

with features x_i*, the true label* $\operatorname{sgn}\{g_1(x_i) - g_2(x_i)\}$*, and the sample weight* $|g_1(x_i) - g_2(x_i)|$*, for subject* i*,* $i = 1, \dots, n$.

Proof of Proposition 1. We observe that to maximize the objective function in (2) is equivalent to maximizing

$$\begin{aligned}
&E_X\big[G_{S|X}\{E(Y|X,S,A=d(X))\}|X\big]\\
&= E_X\big[G_{S|X}\{E(Y|X,S,A=1)\}\mathbb{1}(d(X)=1)\\
&\qquad + G_{S|X}\{E(Y|X,S,A=-1)\mathbb{1}(d(X)=-1)\}\big]\\
&= E_X\big\{\mathbb{1}(d(X)=1)[G_{S|X}\{E(Y|X,S,A=1)\} - G_{S|X}\{E(Y|X,S,A=-1)\}]\\
&\qquad + G_{S|X}\{E(Y|X,S,A=-1)\}\big\}\\
&\propto E_X\big\{\mathbb{1}(d(X)=1)[G_{S|X}\{E(Y|X,S,A=1)\} - G_{S|X}\{E(Y|X,S,A=-1)\}]\big\}.
\end{aligned}$$

□

Proof of Proposition 2. Let $d(x) = \operatorname{sgn}\{f(x)\}$, by this transformation, we consider the following objective on a smooth function $f(x)$,

$$\begin{aligned}
&\arg\max_{d\in\mathbb{D}}\tfrac{1}{n}\textstyle\sum_{i=1}^n\big\{\mathbb{1}(d(x_i)=1)[g_1(x_i)-g_2(x_i)]\big\}\\
&= \arg\max_f \tfrac{1}{n}\textstyle\sum_{i=1}^n \mathbb{1}[\operatorname{sgn}\{f(x_i)\}=1]\cdot[g_1(x_i)-g_2(x_i)]\\
&= \arg\min_f \tfrac{1}{n}\textstyle\sum_{i=1}^n \mathbb{1}\{1\cdot f(x_i)<0\}\cdot[g_1(x_i)-g_2(x_i)]\\
&= \arg\min_f \tfrac{1}{n}\textstyle\sum_{i=1}^n \mathbb{1}[\operatorname{sgn}\{g_1(x_i)-g_2(x_i)\}\cdot f(x_i)<0]\cdot|g_1(x_i)-g_2(x_i)|.
\end{aligned}$$

The sign of the estimated f above is a d to (3).

Hence, the proposed classification-based objective is to minimize

$$\tfrac{1}{n}\textstyle\sum_{i=1}^n \mathbb{1}[\operatorname{sgn}\{g_1(x_i)-g_2(x_i)\}\cdot f(x_i)<0]\cdot|g_1(x_i)-g_2(x_i)|.$$

To this point, we have transformed the optimization problem (2) into a weighted classification problem where for subject i with features x_i, the true label is $\operatorname{sgn}\{g_1(x_i) - g_2(x_i)\}$ and the sample weight is $|g_1(x_i) - g_2(x_i)|$. □

D. Detailed Description on the Algorithm

Algorithm 1 provides an algorithmic overview. The inner expectation $E(Y|X,S,A)$ can be modeled as $\hat{Y}(X,S,A)$ using a twin model separated by the treatment and control groups. For a continuous S, we propose to estimate $G(X,A) = Q_{S|X,A}\{E(Y|X,S,A)\}$ via a quantile regression of $\hat{Y}$ on X but without S. For a discrete S, we propose to obtain an estimate of $G(X,A) = \inf_S\{E(Y|X,S,A)\}$ by finding the minimum among $\{E(Y|X,S=1,A),\ldots,E(Y|X,S=K,A)\}$. The estimated decision rule can then be obtained from the weighted classification. In our implementation, neural networks are used to fit models in the training data sets. A Python package `RISE` based on neural networks is built. Note that the model choices are flexible.

Algorithm 1 RISE (Robust individualized decision learning with sensitive variables)

Input Training data $\mathcal{D}_n = \{Y_i, A_i, X_i, S_i\}_{i=1}^n$
Output Estimated decision rule $\hat{d}$

1: $\hat{Y}_i(x_i, s_i, a_i) \leftarrow$ Model $E(Y|X,S,A=a)$ using $\mathcal{D}_n$ with $a=1$ and $a=-1$, respectively.
2: **if** S is continuous **then**
3: $\quad g_1(x_i) \leftarrow$ Model $Q_{S|X,A}\{E(Y|X,S,A=a)\}$ via quantile regressions of $\hat{Y}_i(x_i,s_i,a_i)$ on x_i, for $\mathcal{D}_n$ with $a=1$.
4: $\quad g_2(x_i) \leftarrow$ Model $Q_{S|X,A}\{E(Y|X,S,A=a)\}$ via quantile regressions of $\hat{Y}_i(x_i,s_i,a_i)$ on x_i, for $\mathcal{D}_n$ with $a=-1$.
5: **if** S is discrete **then**
6: $\quad g_1(x_i) \leftarrow$ Compute $\inf_{s\in\mathbb{S}}\{\hat{Y}_i(x_i,s,a_i=1)\}, \forall i$.
7: $\quad g_2(x_i) \leftarrow$ Compute $\inf_{s\in\mathbb{S}}\{\hat{Y}_i(x_i,s,a_i=-1)\}, \forall i$.
8: $\hat{d} \leftarrow$ Build a weighted classification model with features x_i, label $\mathrm{sgn}\{g_1(x_i)-g_2(x_i)\}$, and sample weight $|g_1(x_i)-g_2(x_i)|$.
9: **Return** $\hat{d}$

D.1. Details on Modeling and Hyperparameter Tuning

In our implementation, neural networks with mean or quantile losses are used to fit the models with hyperparameters tuned via a 5-fold cross validation in the training data sets. Specifically, implemented in TensorFlow (Abadi et al., 2016), neural networks with mean squared loss is used to model $E(Y|X,S,A)$ separated by the control arm and the treatment arm, respectively. For continuous S, to model $Q_{S|X,A}\{E(Y|X,S,A)\}$, neural networks with quantile loss is used with a prespecified τ, for the control arm and the treatment arm, respectively. In the final weighted classification model, neural networks with cross-entropy loss is used. Note that the model choices here are flexible. One can perform model selection if they would like to.

Hyperparameter tuning helps prevent overfitting and is essential in machine learning methods or other black-box algorithms such as neural networks. In our implementation, the optimal hyperparameters are obtained via a 5-fold cross validation in the training data sets. Specifically, we consider the number of hidden layers (1, 2, and 3 layers), the number of hidden units in each layer (256, 512, and 1024 nodes), activation function (RELU, Sigmoid, and Tanh), optimizer (Adam, Nadam, and Adadelta), dropout rate (0.1, 0.2, and 0.3), number of epochs (50, 100, and 200), and batch size (32, 64, and 128).

E. Compared Approaches

For comparison, we consider the naive and mean-optimal approaches described in Section 2.1, which correspond to different choices of $G(\cdot)$ functions. The naive decision rule that simply disregard information of S, denoted as **Base**, can be formulated in our optimization framework of (2) by letting $G(X,A) = E(Y|X,A)$. The IDR can be estimated directly by fitting a model of Y on X in each treatment arm. The resulting IDR is not sensitive variables-aware and is biased due to confounding, as discussed. Another IDR that resembles traditional mean-optimal decision rules, denoted as **Exp**, can be formulated as $G(X,S,A) = E(Y|X,S,A)$. This can be obtained by training a classification model without S, i.e., only using X, after obtaining an outcome model for the inner expectation $E(Y|X,S,A)$. Note that this approach is not robust to extreme behaviors in S. The modeling approaches described in Appendix D applies to here. We also remark that there is limited work in the IDR-related literature that can be immediately compared with ours.

F. Simulation Details

For simulation, we consider training data and the testing data, respectively, with sample sizes of 5,000. All results are based on 100 replications.

Example 1. Here we provide the detail for the simulation of the motivating example introduced in Section 1. The outcome is generated using the following model: $Y_i = \mathbb{1}(X_i > 0.5)\{5 + 10\mathbb{1}(A_i = 1) + 22S_i - 24\mathbb{1}(A_i = 1)S_i\} + \mathbb{1}(X_i \leq 0.5)\{11 + 19\mathbb{1}(A_i = 1) + 2S_i - 32\mathbb{1}(A_i = 1)S_i\} + \epsilon_i$, where the covariate $X_i \sim Unif[0, 1]$, treatment assignment $A_i \sim Bernoulli(0.5)$, and the noise $\epsilon_i \sim N(0, 1)$. For a discrete type S, the sensitive variable $S_i \sim Bernoulli(0.5)$. For a continuous type S, S_i is generated from a mixture of beta distributions, $Beta(4, 1)$ and $Beta(1, 4)$, with equal mixing proportions.

Example 2. We generate the outcome Y using the following model: $Y_i = \{0.5 + \mathbb{1}(A_i = 1) + \exp(S_i) - 2.5S_i\mathbb{1}(A_i = 1)\}\{1 + X_{i1} - X_{i2} + X_{i3}^2 + \exp(X_{i4})\} + \{1 + 2\mathbb{1}(A_i = 1) + 0.2\exp(S_i) - 3.5S_i\mathbb{1}(A_i = 1)\}\{1 + 5X_{i1} - 2X_{i2} + 3X_{i3} + 2\exp(X_{i4})\} + \epsilon$, where $X_{ij} \sim U(0, 1),\ j = 1, \dots, 6,\ A$ satisfies $\log\{P(A_i = 1|X_i)/P(A = 0|X_i)\} = -0.6(S_i + X_{i1} - X_{i2} + X_{i3} - X_{i4} + X_{i5} - X_{i6})$, and $\epsilon_i \sim N(0, 1)$. For a continuous type S, S_i is generated from a mixture of beta distributions, $Beta(4, 1)$ and $Beta(1, 4)$, with equal mixing proportions; for a discrete type S, we consider a binary S_i that satisfies $\log\{P(S_i = 1|X_i)/P(S_i = 0|X_i)\} = -2.5 + 0.8(X_{i1} + X_{i2} + X_{i3} + X_{i4} + X_{i5} + X_{i6})$.

Table 3 summarizes the performance of the proposed IDRs compared to the mean criterion for Example 1 and Example 2. The proposed RISE achieves the largest objectives and improves the value among vulnerable subjects, while maintaining comparative overall values. As demonstrated in the toy example introduced in Section 1, we expect that RISE helps improve the value among the vulnerable subjects while maintaining a comparable overall value. As for the objective, intuitively, the proposed rule is expected to have a larger objective. We also point out that there is no direct relationship between the objective among all subjects versus the objective among vulnerable subjects. For example, using the toy example with setup in Table 1, and limiting to subjects with $X \leq 0.5$ only, $S = 1$ is vulnerable and is assigned $A = -1$ by the proposed RISE. The objective among $S = 1$ is 13 but the objective among both $S = 0$ and $S = 1$ is $12 = (11 + 13)/2$, which is smaller than that among the vulnerable group. In other words, by protecting the vulnerable subjects, the proposed rule may lead to an increase in the outcome of the vulnerable group, and the gain may result in a higher outcome than the overall mean outcome.

Table 3: Simulation results for Example 1 and Example 2. Standard error in parenthesis.

Example	Type of S	IDR	Obj. (all)	Obj. (vulnerable)	Value (all)	Value (vulnerable)
1	Disc.	Base	7.03 (0.03)	7.01 (0.04)	14.3 (0.05)	7.92 (0.06)
		Exp	6.39 (0.03)	6.39 (0.04)	**14.4** (0.05)	7.14 (0.06)
		RISE	**12.0** (0.01)	**12.0** (0.01)	13.0 (0.01)	**14.0** (0.01)
	Cont.	Base	9.12 (0.03)	9.14 (0.04)	14.5 (0.08)	8.25 (0.11)
		Exp	8.75 (0.03)	8.75 (0.04)	**14.6** (0.08)	7.58 (0.06)
		RISE	**10.3** (0.02)	**10.3** (0.03)	13.9 (0.04)	**10.3** (0.06)
2	Disc.	Base	7.79 (0.02)	8.66 (0.03)	19.4 (0.04)	11.4 (0.06)
		Exp	9.12 (0.03)	10.1 (0.03)	**19.5** (0.04)	14.4 (0.05)
		RISE	**13.5** (0.01)	**14.0** (0.01)	17.4 (0.02)	**22.1** (0.02)
	Cont.	Base	9.89 (0.02)	9.87 (0.03)	17.6 (0.02)	9.09 (0.04)
		Exp	11.1 (0.02)	11.1 (0.02)	**17.8** (0.02)	12.2 (0.04)
		RISE	**13.7** (0.02)	**13.7** (0.02)	17.0 (0.01)	**18.9** (0.03)

F.1. Additional Simulations

Different quantile criteria. For the quantile criteria, we also consider τ = 0.1 and 0.5, respectively. Table 4 presents the simulation results for Example 2 with continuous S using 0.1 quantile criterion and 0.5 quantile criterion, respectively.

S as a noise variable. We generate the outcome Y using the following model where S is not involved: $Y = \mathbb{1}(X_1 \leq 0.5)\{8+12\mathbb{1}(A = 1)+16\exp(X_2)-26\mathbb{1}(A = 1)X_2\}+\mathbb{1}(X_1 > 0.5)\{13+3\mathbb{1}(A_i = 1)+2\exp(X_2)-8\mathbb{1}(A = 1)X_2\}+\epsilon$, where $X_j \sim U(0, 1),\ j = 1, 2,\ A \sim Bernoulli(0.5)$, and $\epsilon \sim N(0, 1)$. For continuous S, $S = \text{expit}\{-2.5(1-X_1-X_2)\}$; for discrete S, we consider a binary S that satisfies $\log\{P(S = 1|X)/P(S = 0|X)\} = -2.5(1 - X_1 - X_2)$. Table 5 summarizes the performance of the proposed IDRs compared to the mean criterion for Example 2. The estimated objective and value function are similar for the compared IDRs, which indicates the robustness of the proposed RISE.

Table 4: Simulation results for Example 2 with continuous S using 0.1 quantile criterion and 0.5 quantile criterion, respectively. Standard error in parenthesis.

Type of S	τ	IDR	Obj. (all)	Obj. (vulnerable)	Value (all)	Value (vulnerable)
Cont.	0.1	Base	7.93 (0.03)	7.92 (0.03)	17.7 (0.02)	8.64 (0.07)
		Exp	8.88 (0.05)	8.85 (0.05)	**17.8** (0.02)	10.6 (0.12)
		RISE	**13.8** (0.01)	**13.7** (0.02)	16.9 (0.01)	**20.9** (0.03)
Cont.	0.5	Base	17.3 (0.04)	17.2 (0.04)	17.7 (0.02)	23.8 (0.19)
		Exp	17.2 (0.03)	**17.4** (0.03)	**17.8** (0.02)	22.1 (0.17)
		RISE	**17.4** (0.04)	**17.4** (0.04)	**17.8** (0.02)	**24.0** (0.22)

Table 5: Simulation results for scenario when S is a noise variable. Vulnerable subjects cannot be defined as S is not important in the example. The estimated objective and value function are similar for the compared IDRs, which indicates the robustness of the proposed RISE.

Type of S	IDR	Obj. (all)	Obj. (vulnerable)	Value (all)	Value (vulnerable)
Disc.	Base	27.5 (0.03)	-	27.5 (0.06)	-
	Exp	27.5 (0.03)	-	27.5 (0.06)	-
	RISE	27.5 (0.03)	-	27.5 (0.06)	-
Cont.	Base	27.2 (0.04)	-	27.3 (0.07)	-
	Exp	27.2 (0.04)	-	27.3 (0.07)	-
	RISE	27.2 (0.04)	-	27.3 (0.07)	-

G. Detailed Information and Results for Real-data Applications

For real-data applications, we consider a 80-20 split of the dataset into a training data and a testing data. Continuous covariates are standardized before the estimation. All results are based on 100 replications.

Fairness in a job training program. To illustrate the implication of the proposed method from a fairness perspective, we consider the National Supported Work (NSW) program (LaLonde, 1986) for improving personalized recommendations of a job training program on increasing incomes. This program intended to provide a 6 to 18-month training for individuals in face of economic and social problems such as former drug addicts and juvenile delinquents. The original experimental dataset consists of 185 individuals who received the job training program ($A = 1$) and 260 individuals who did not ($A = -1$). The baseline covariates are age, years of schooling, race (1 = African Americans or Hispanics, 0 = others), married (1 = yes, 0 = no), high school diploma (1 = yes, 0 = no), earning in 1974, and earning in 1975. The outcome variable is the earning in 1978. In the exploratory analysis using causal forest (Wager & Athey, 2018), we observe that age may play an important role in the causal effect of the job training program on the long-term post-market earning. In the following data example we use age as the sensitive variable S and other baseline covariates as X. The earnings in years 1974, 1975, and 1978 are transformed by taking the logarithm of the earning plus one.

Improvement of HIV treatment. To illustrate the implication of the proposed method from a safety perspective when there is delayed information, we consider the ACTG175 dataset among HIV positive patients (Hammer et al., 1996). The original study considers a total of 2,139 patients who were randomly assigned into four treatment groups. In this data application, we focus on finding the optimal IDRs between two treatments: zidovudine combined with didanosine ($A = -1$) and zidovudine combined with zalcitabine ($A = 1$). The total number of patients receiving these two treatments is 1,046. The baseline covariates we consider are age, weight, CD4 T-cell amount at baseline, hemophilia (1 = yes, 0 = no), homosexual activity (1 = yes, 0 = no), Karnofsky score, history of intravenous drug use (1 = yes, 0 = no), gender (1 = male, 0 = female), CD8 T-cell amount at baseline, race (1 = non-Caucasian, 0 = Caucasian), number of days of previously received antiretroviral therapy, use of zidovudine in the 30 days prior to treatment initiation (1 = yes, 0 = no), and symptomatic indicator (1 = symptomatic, 0 = asymptomatic). The outcome variable is the CD4 T-cell amount at 96 ± 5 weeks from the baseline. We consider CD8 T-cell amount at baseline as the sensitive variable. The response of CD8 T-cell among HIV positive patients has not been fully understood (Boppana & Goepfert, 2018). Clinically, it is plausible that only CD4 is measured in clinical visits where treatments are based on, hence CD8 might not be measured and not used in decision making. As our exploratory analysis using causal forest shows, CD8 T-cell amount may play an important part in the treatment effect of the outcome.

Safe resuscitation for patients with sepsis. For this application, we apply the proposed method to treating sepsis, a life-threatening disease. This application intends to provide an example to apply our method with multiple categorical sensitive variables in the scenario where there is missing yet important information at the time of decision making. We apply the proposed method to a sepsis study from the University of Pittsburgh Medical Center (UPMC). The original study cohort includes 30,687 patients with Sepsis-3 (Seymour et al., 2016) within 6 hours of hospital arrival from 14 UPMC hospitals between 2013 and 2017. For our data analysis, we consider X to be baseline patient characteristics 4 hours before sepsis onset, which includes patient demographics of age, gender (1 = male, 0 = female), race (1 = Caucasian, 0 = others), and weight, and vital signs of usage of mechanical ventilation (1 = yes, 0 = no), respiratory rate, temperature, intravenous fluids (1 = yes, 0 = no), Glasgow Coma Scale score, platelets, blood urea nitrogen, white blood cell counts, glucose, creatinine. We consider two sensitive variables, lactate and Sequential Organ Failure Assessment (SOFA) score 4 hours before sepsis onset. Note that their measurements are obtained retrospectively after treatment decisions have been made and are not available at times of decision. According to the new definition of Sepsis-3 (Shankar-Hari et al., 2016), a serum lactate level >2 mmol/L is considered to be in critical conditions and is highly likely to indicate a septic shock. Also, a SOFA score greater than 6 has been associated with a higher mortality (Vincent et al., 1996; Ferreira et al., 2001). The treatment option is whether the patient took any vasopressors during the first 24 hours after sepsis onset. The outcome is patient survival at day 90. The analysis cohort contains 6,539 patients in total. We are interested in making decision about whether to treat patients with vasopressors in the first 24 hours after sepsis onset given the measurements of lactate and SOFA are not available at the time of decision making.

Results. Table 6 presents the performance of various IDRs on the three applications. As expected, RISE has the largest objective as well as value among vulnerable subjects. The patterns are similar to that in the synthetic experiments in Section 3.1. In applications to the job training data and the sepsis study, results show that RISE has a larger value among all subjects than other IDRs. This is possible when there are more gains in the vulnerable subjects than other subjects, which further demonstrate the superiority of the proposed approach in improving worst-case outcomes caused by sensitive variables. We provide visualizations by Shapley additive explanations (SHAP) (Lundberg & Lee, 2017) for RISE and Exp, respectively, in Appendix G about feature importance in the final classification models to help interpret important covariates in making the decisions. The SHAP approach provides united values to describe the correlation between each feature and the predicted decision rule, respectively (Lundberg & Lee, 2017). Overall, the direction of correlations is similar for RISE and Exp, but their ranking of feature importance may be different.

Table 6: Estimated objective and value of different IDRs for the three data applications. Standard error in parenthesis. The outcome of each study is italicized.

Dataset	IDR	Obj. (all)	Obj. (vulnerable)	Value (all)	Value (vulnerable)
NSW *log(income+1)*	Base	5.26 (0.04)	5.28 (0.05)	6.32 (0.05)	6.33 (0.07)
	Exp	5.22 (0.04)	5.24 (0.05)	6.37 (0.05)	6.37 (0.07)
	RISE	**5.43** (0.04)	**5.44** (0.04)	**6.42** (0.04)	**6.42** (0.06)
ACTG175 *CD4 T-cell amount*	Base	336.9 (1.65)	338.1 (2.23)	350.5 (1.86)	357.5 (2.24)
	Exp	337.5 (1.65)	338.9 (1.80)	**351.9** (1.95)	359.1 (2.21)
	RISE	**351.5** (1.67)	**351.2** (1.80)	351.8 (1.88)	**363.1** (2.19)
Sepsis *survival rate*	Base	0.752 (0.001)	0.721 (0.001)	0.965 (0.001)	0.905 (0.002)
	Exp	0.752 (0.001)	0.721 (0.002)	0.966 (0.001)	0.908 (0.002)
	RISE	**0.771** (0.001)	**0.735** (0.001)	**0.972** (0.001)	**0.923** (0.002)

G.1. Additional Background on the Sepsis Application

Sepsis is leading cause of acute hospital mortality and commonly results in multi-organ dysfunction among ICU patients (Sakr et al., 2018). Clinically, treatment decisions for sepsis patients are needed to be made within a short period of time due to the rapid deterioration of patient conditions. Lactate and the Sequential Organ Failure Assessment (SOFA) score have been two important indicators of sepsis severity and has been found to be more useful for predicting the outcome of sepsis than other clinical vitals and comorbidity scores (Howell et al., 2007; Krishna et al., 2009; Shankar-Hari et al., 2016). Typically, information of baseline patient characteristics such as age, gender, race, and weight, and common vital signs such as usage of mechanical ventilation, respiratory rate, temperature, intravenous fluids, Glasgow Coma Scale score, platelets, blood urea nitrogen, white blood cell counts, glucose, and creatinine are obtained at the admission of patients. On the other hand, SOFA score combines performance of several organ systems in the body such as neurologic, blood, liver, and kidney (Seymour et al., 2016) and cannot be obtained directly. Lactate labs measures the level of lactic acid in the blood (Andersen et al., 2013) and are less common in routine examination, which could be delayed in ordering. Hence, their information may not be available by the time of treatment decision due to multiple reasons including doctors' delayed ordering, long laboratory processing time, or the rapid deterioration of development of sepsis, which poses tremendous difficulties for early diagnosis and treatment decisions within a short time.

G.2. Data Availability

The access of the job training dataset (LaLonde, 1986) is available at `https://users.nber.org/~rdehejia/data/.nswdata2.html`. The access of the ACTG175 dataset (Hammer et al., 1996) is available from the R package `speff2trial`. The third dataset (Seymour et al., 2016) is not publicly available. All real data used in the paper are deidentified with no personal information.

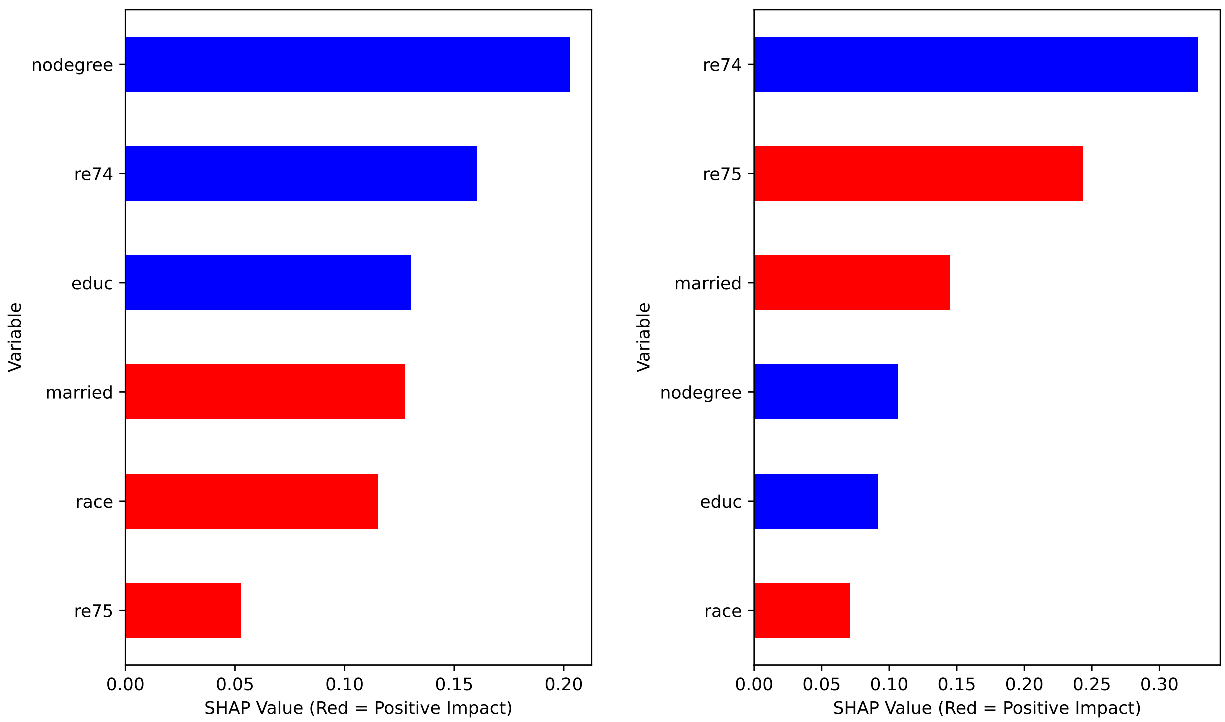
G.3. Visualization

Here we provide visualizations of features that are important in the estimated decision rules for the three real-data applications in Section 3.2. The Shapely additive explanations (SHAP) (Lundberg & Lee, 2017) is considered to be a united approach to explaining the predictions of any machine learning or black-box models. Figure 1, Figure 2, and Figure 3 presents the SHAP variable importance plots in the final weighted classification model by RISE and Exp, respectively, for the three real-data applications. Correlations between the feature and their SHAP value are highlighted in color. The red color means a feature is positively correlated with assigning treatment A = 1 and the blue indicates a negative correlation. Overall, the direction of correlation is similar for RISE and Exp, but their ranking of feature importance may be different.

Fairness in a job training program. Figure 1 presents the SHAP variable importance plots in the final weighted classification model by RISE and Exp, respectively. We observe that whether having a high school diploma and income in 1974 are two important features in the variable important plot by RISE, while incomes in 1974 and in 1975 are important by Exp. It seems that being no degree and low income in 1974 has a higher chance of assigning $A = 1$ (to receive the job training program) by RISE, while low income in 1974 and but a higher income in 1975 may be associated with assigning $A = 1$ by Exp.

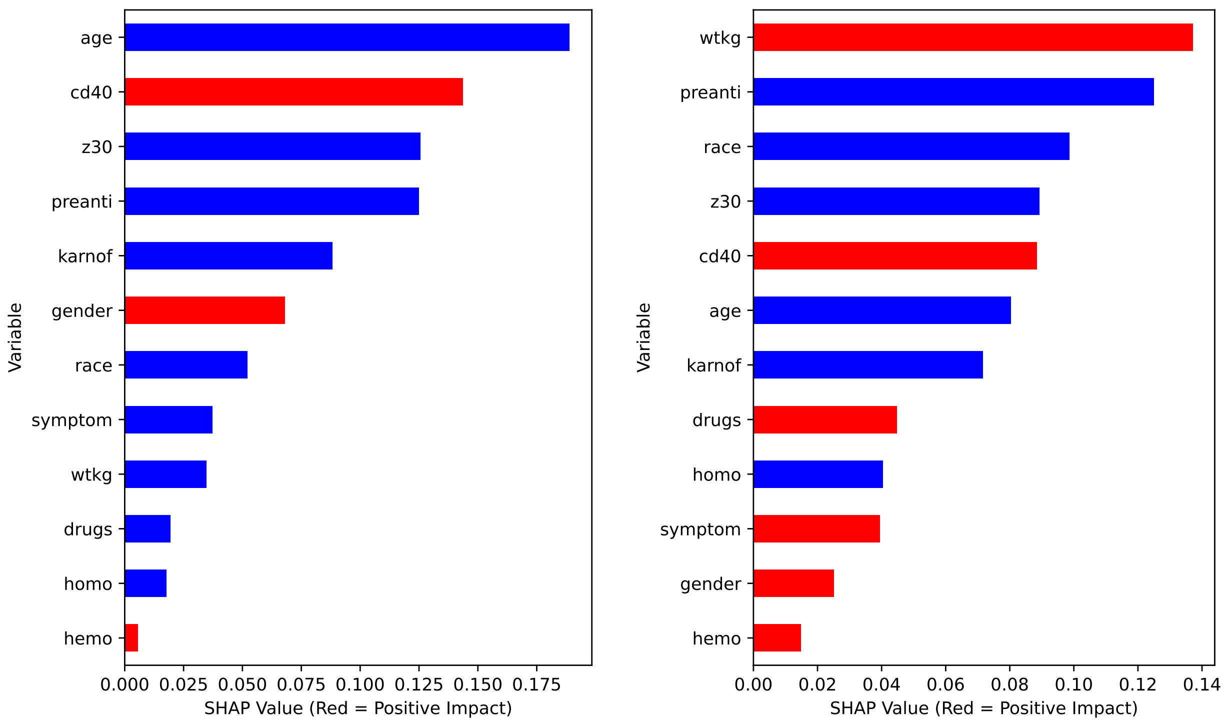
Improvement of HIV treatment. Figure 2 presents the SHAP variable importance plots in the final weighted classification model by RISE and Exp, respectively. We observe that age and CD4 T-cell counts are two important features in the variable important plot by RISE, while weight and number of days of previously received antiretroviral therapy are important by Exp. It seems that being of a younger age and high CD4 T-cell count has a higher chance of assigning $A = 1$ (zidovudine combined with didanosine) by RISE, while being of a larger weight and few days of previously received antiretroviral therapy may be associated with assigning the treatment by Exp.

Safe resuscitation for patients with sepsis. Figure 3 presents the SHAP variable importance plots in the final weighted classification model by RISE and Exp, respectively. We observe that Glasgow Coma Scale score, age, and platelets appears to be important features in both the plot by RISE and that by Exp. Other important features in the plot by RISE include temperature and blood urea nitrogen, where in the plot by Exp, respiratory rate and white blood cell counts are of top importance. Being in a low temperature with a high blood urea nitrogen tends to be predicted as $A = 1$ (to assign vasopressors) by RISE while being of higher respiratory rate with high white blood cell counts tends to be predicted as $A = 1$ by Exp.



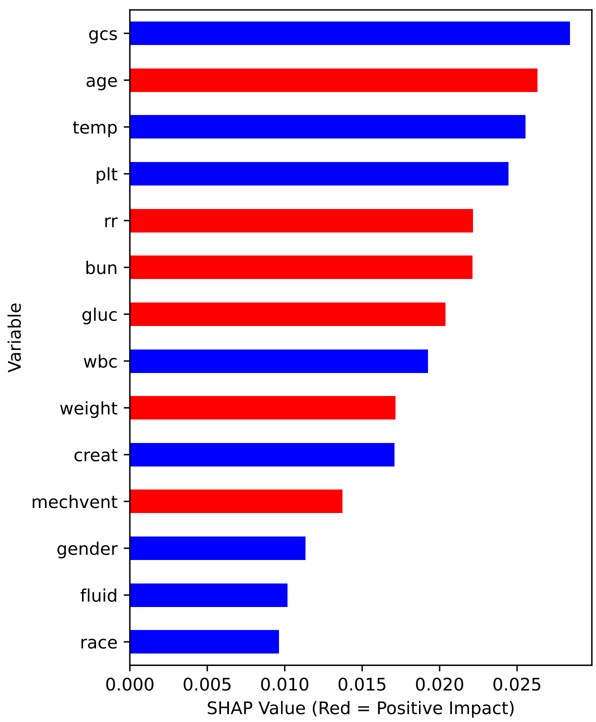
(a) (b)

Figure 1: Visualization for the job training program: SHAP variable importance plots for decision rules RISE (a) and Exp (b), respectively. Covariates (X) are ranked by variable importance in descending order. Correlations between the feature and their SHAP value are highlighted in color. The red color means a feature is positively correlated with assigning treatment $A = 1$ and the blue indicates a negative correlation.

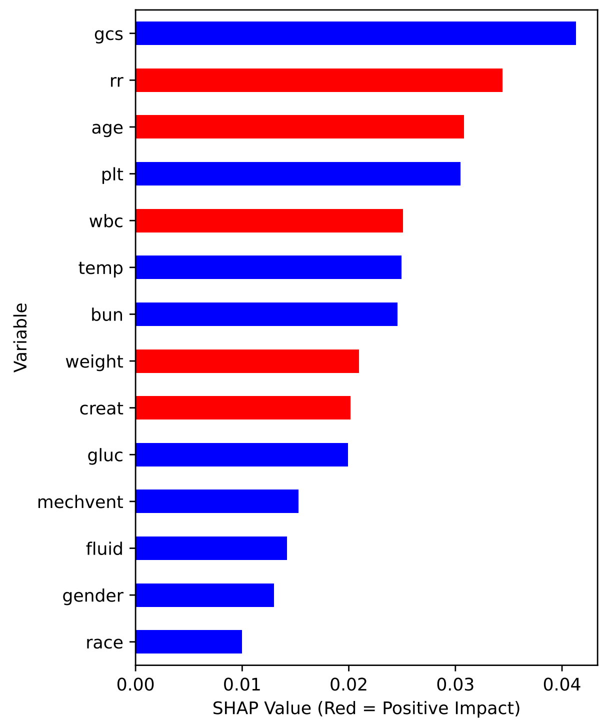


(a) (b)

Figure 2: Visualization for the ACTG175 dataset: SHAP variable importance plots for decision rules RISE (a) and Exp (b), respectively. Covariates (X) are ranked by variable importance in descending order. Correlations between the feature and their SHAP value are highlighted in color. The red color means a feature is positively correlated with assigning treatment $A = 1$ and the blue indicates a negative correlation.



(a)



(b)

Figure 3: Visualization for the sepsis data: SHAP variable importance plots for decision rules RISE (a) and Exp (b), respectively. Covariates (X) are ranked by variable importance in descending order. Correlations between the feature and their SHAP value are highlighted in color. The red color means a feature is positively correlated with assigning treatment $A = 1$ and the blue indicates a negative correlation.