Active Learning for Improving Decision-Making from Imbalanced Data

Iiris Sundin[1] **Peter Schulam**[*2] **Eero Siivola**[*1] **Aki Vehtari**[1] **Suchi Saria**[2] **Samuel Kaski**[1]

Abstract

This extended abstract considers the reliability of prediction-based decision-making in a task of deciding which action a to take for a target unit after observing its covariates $\tilde{x}$ and predicted outcomes $\hat{p}(\tilde{y} \mid \tilde{x}, a)$. An example case is personalized medicine and the decision of which treatment to give to a patient. A common problem when learning these models from observational data is imbalance, that is, difference in treated/control covariate distributions. We propose to assess the decision-making reliability by estimating the model's Type S error rate, which is the probability of the model inferring the sign of the treatment effect wrong. Furthermore, we use the estimated reliability as a criterion for active learning, in order to collect new (possibly expensive) observations, instead of making a forced choice based on unreliable predictions. We demonstrate the effectiveness of this decision-making aware active learning in two decision-making tasks: in simulated data with binary outcomes and in a medical dataset with synthetic and continuous treatment outcomes.

1. Introduction

In this extended abstract, we consider a case where machine learning is used to provide predictions of outcomes under alternative actions, to aid human decision-making (Schulam & Saria, 2017). To fit a model to this task, we need data recording previous actions a, observed outcomes y, and any features relevant to the context of the decision, x. Then, the goal is to estimate $p(Y \mid X = x, A = a)$ and, further, individual treatment effect (ITE), $\tau(x) = \mathbb{E}[Y[1] - Y[0] \mid X = x]$, where $Y[a]$ denotes the potential outcome of treatment $A = a$ (Rubin, 1978). ITE provides sufficient information to choose between two actions.

*Equal contribution [1]Department of Computer Science, Aalto University, Espoo, Finland [2]Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218, USA. Correspondence to: Iiris Sundin <iiris.sundin@aalto.fi>.

Real-world Sequential Decision Making workshop at ICML 2019.

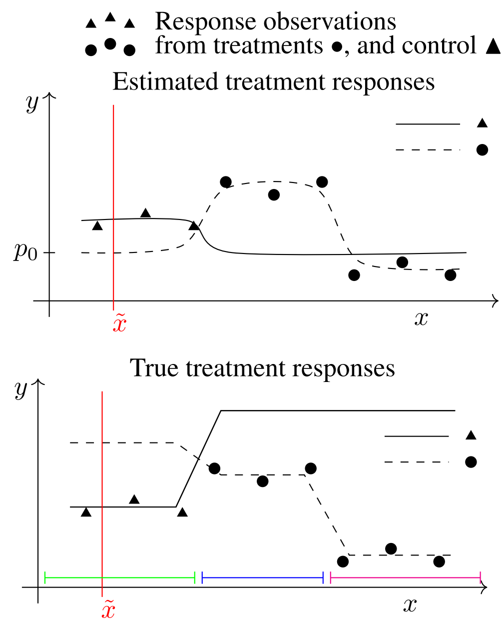


Figure 1. An example decision-making task is to choose a treatment for a specific $\tilde{x}$ (red mark). Upper graph shows the posterior means of the potential outcome models given observations, and the prior mean p_0. The true responses are in the lower graph, which shows that there are three regions with different response types (marked in green, blue and magenta). Lower y is better. The problem is that there are no observations about treatment ● in the green region, which causes it to appear to be the best choice for $\tilde{x}$, although that is incorrect.

The estimation of ITE is susceptible to many error sources (Pearl, 2009; Schulam & Saria, 2017; Mitchell et al., 2018), of which we concentrate on imbalance (Gelman & Hill, 2007).

Imbalance is defined as the difference in covariate distributions in the treated and control groups. Imbalance makes correct model specification essential for avoiding bias in treatment effect estimates (Gelman & Hill, 2007). Mis-specification of the model could be avoided by using non-parametric models, but their variance increases quickly when extrapolating. Recently, imbalance has been shown to increase the upper bound of the model error in estimation of ITE (Shalit et al., 2017). Furthermore, imbalance becomes the more prevalent issue the higher-dimensional the covariate space is (D'Amour et al., 2018).

There are many existing ways to deal with imbalance when learning the average treatment effect (ATE). In causal inference, the most common methods are propensity score matching or weighting (Rosenbaum & Rubin, 1984; Hirano et al., 2003; Lunceford & Davidian, 2004), and modeling the potential outcomes (Imbens & Rubin, 2015; Hernán & Robins, 2018), as well as doubly robust methods which implement both (Bang & Robins, 2005; Funk et al., 2011).

Even though these methods can decrease bias in treatment effect estimates, they will increase variance, and therefore may make the decision-making less reliable. This is especially the case with ITE; For example, in the areas of covariate space where there are more control units than treated, intuition is that the model for the treated outcome either has to generalize from less-representative observations (increasing bias) or extrapolate (increasing variance). Either way, there is higher *uncertainty* about the treated outcome, which makes reliable decision-making difficult. An extreme case of this is illustrated in Fig. 1. A natural question then is, could other data sources be exploited instead of making a forced choice based on insufficient observational data.

This work has three main contributions, which are complementary to each other and can be used independently. First, we describe how imbalance decreases *decision-making performance* by increasing Type S error rate (Gelman & Tuerlinckx, 2000), which is the probability of the model inferring the sign of the treatment effect wrong. Second, we propose a Bayesian *estimate for the Type S error rate*, which allows quantifying the reliability of a decision-support model. Third, we propose to alleviate the consequences of imbalance by actively collecting more data (e.g. from a simulator, new experiment, or by asking a domain expert). To this end, we introduce *decision-making aware active learning criteria* that improve decision-making performance by minimizing the estimated Type S error rate.

Finally, in many cases there are restrictions on what can be measured. For example, in medicine it is in general not ethical to do an experiment on a patient in order to get information to improve the treatment plan of another patient. Therefore, regular active learning would not be possible and, instead, any new information has to be acquired indirectly. For this reason, we introduce the idea of *counterfactual elicitation* which means soliciting indirect observations about counterfactual outcomes, that is, what would have been the outcome had x been treated with a' instead of a. We demonstrate the effectiveness of the proposed active learning criteria applied to counterfactual elicitation.

2. Related Work

Active learning proposed by Javdani et al. (2014) aims at improving automated decision-making by reducing model uncertainty so that the remaining hypotheses are confined to the same decision region. The importance of active learning for decision-making tasks has been noted in other fields, where e.g. Saar-Tsechansky & Provost (2007) developed a heuristic method for deciding which consumers to target in marketing campaigns. Closest to our work is the work on active learning with logged data (Yan et al., 2018), which proposes a de-biasing query strategy for a classification task. The difference to our work is that they assume the propensities (probability of revealing the label) to be known.

Using data to estimate the effect of interventions has been extensively studied in the field of causal analysis (see e.g. Pearl (2009); Morgan & Winship (2014); Imbens & Rubin (2015); Hernán & Robins (2018)). The issue of imbalance is discussed by Johansson et al. (2016), and they propose an approach based on empirical risk minimization and domain shift to improve predictions. Xu et al. (2016) estimate the individual treatment effects in a longitudinal setting where individual-specific treatment parameters are refined over time as more observations are collected. Alaa & van der Schaar (2017) use Gaussian processes to model individual-specific outcomes under alternative treatments and prove minimax rates on the risk that the approach achieves. Our work builds on these ideas, and is most closely related to the works by Xu et al. (2016) and Alaa & van der Schaar (2017). Although we also use Gaussian processes to model treatment effects, our work is unique in that it leverages the probabilistic framework to design an active learning algorithm to improve the decision-making process.

3. Methods

Preliminaries. Causal effect of a treatment is the difference between outcomes when a unit $\mathbf{x} \in \mathcal{X}$ is treated and not treated, where $\mathcal{X}$ is the population. Individualized treatment effect is defined as $\tau(\mathbf{x}) = \mathbb{E}[Y[1] - Y[0] \mid X = \mathbf{x}]$, where $a = 1$ means treated and $a = 0$ not treated, i.e. control. The estimated treatment effect is $\hat{\tau}(\mathbf{x})$.

Setup. The outcome model has been learned in retrospective from observational data, which may be *imbalanced* in the observed actions. The observational data D is a set of n observations $\{y_i, a_i, \mathbf{x}_i\}_{i=1}^{n}$, where y_i is the observed outcome of action a_i for unit $\mathbf{x}_i$. Imbalance means that the covariate distributions are different in the treated ($a_i = 1$) and control groups ($a_i = 0$).

We assume that the unknown policy used to choose actions in the training data only depends on the observed covariates $\mathbf{x} \in \mathcal{X}$. This is equivalent to the no unmeasured confounders assumption (Hernán & Robins, 2018) and implies that all confounders are included in $\mathbf{x}$. We further assume consistency of potential outcomes, which means that the potential outcomes $p(y[a] \mid \mathbf{x}) = p(y \mid X = \mathbf{x}, A = a)$ can be directly estimated from the observed outcomes in the training data. Regardless of this, imbalance will still cause issues.

In this work, the task is to decide which action a to choose for $\tilde{\mathbf{x}}$. Decision-making performance is measured as the probability of correct decision, or equivalently, the proportion of correct decisions in repeated decision-making tasks.

Definition: Type S error rate γ is the probability of the model inferring the sign of the treatment effect wrong;
γ: $\mathcal{M} \times P_{\mathcal{X},\mathcal{Y}} \to [0, 1]$ where $\mathcal{M}$ is the model space and $P_{\mathcal{X},\mathcal{Y}}$ is the true treatment effects (distributions over $\mathcal{X} \times \mathcal{Y}$). The expected Type S error rate in $\mathcal{X}$ is
$\gamma = \mathbb{E}_{P_{\mathcal{X},\mathcal{Y}}}[\mathbb{I}(\text{sign}(\hat{\tau}) \neq \text{sign}(\tau))]$, where $\mathbb{I}(A) = 1$ if condition A is true, and 0 otherwise.

The expected proportion of correct decisions in population $\mathcal{X}$ is then $1 - \gamma$, which makes Type S error rate a natural measure of the decision-making performance.

We assume that the decision-maker's policy is to choose the action with the highest expected utility of the outcome. Without loss of generality we will assume that the utility is directly the outcome y (higher better).

3.1. Effect of Imbalance on Type S Error Rate

In this section we give intuition on how imbalance increases the error rate in decision-making. First, assume a probabilistic model of potential outcomes, with broad prior distributions. This implies that when the sample size is small, posteriors will be wide. Then, under certain assumptions, it can be shown that imbalance decreases the expected number of samples locally, which makes posteriors wider and therefore increases the Type S error rate locally.

Then the question remains whether the error rate increases globally as well, or do the local effects cancel out each other. The intuition is that since the error rate increases exponentially with decreasing number of samples, if there is high-enough density in areas with high imbalance, then the local increase in the error rate cannot be compensated elsewhere.

3.2. Estimated Type S Error Rate

In the Bayesian sense, a model $\hat{p}(y[a] \mid \mathbf{x}, D)$ captures our current understanding of the problem, and therefore the estimated Type S error rate is $\hat{p}(y[1] < y[0] \mid \mathbf{x}, D)$ if the expected effect is positive, that is, if $\mathbb{E}_{\hat{p}(y[1]|\mathbf{x},D)}[y[1]] > \mathbb{E}_{\hat{p}(y[0]|\mathbf{x},D)}[y[0]]$. Respectively, if the expected effect is negative, then the estimated error rate is $\hat{p}(y[1] > y[0] \mid \mathbf{x}, D)$.

For example, when the predicted treatment effect $\tau = y[1] - y[0]$ is normally distributed, then it is easy to show that the estimated Type S error rate in a test unit $\tilde{\mathbf{x}}$ is the tail-probability

$$\hat{\gamma}(\tilde{\mathbf{x}}) = \text{probit}^{-1}\left(-\frac{|\mathbb{E}_{\hat{p}(\tilde{\tau}|\tilde{\mathbf{x}},D)}[\tilde{\tau}]|}{\text{Var}(\hat{p}(\tilde{\tau} \mid \tilde{\mathbf{x}}, D))^{\frac{1}{2}}}\right), \tag{1}$$

where probit^{-1} is the cumulative distribution function of normal distribution, and the expectations and variances are over the posterior predictive distribution of $\tilde{\tau} \mid \tilde{\mathbf{x}}$. From (1) we see that the estimated Type S error rate will increase if the estimated treatment effect decreases, or if posterior uncertainty (variance) increases. Intuitively this makes sense.

3.3. Decision-Making Aware Active Learning

Our hypothesis is that active learning criteria that reduce the estimated Type S error rate will result in higher decision-making performance. We call active learning criteria that reduce the estimated Type S error *decision-making aware*.

We propose to minimize the estimated Type S error rate in eq. (1) with active learning. The criterion is to maximize the estimated reliability of a decision at $\tilde{x}$, that is, $1 - \hat{\gamma}(\tilde{x})$. Directly minimizing the error can be interpreted as only exploiting what we already know, which lacks in exploration. There is an easy fix, though, which is to add exploration on the error. The exploration-exploitation trade-off is managed by maximizing the expected information gain on the predictive distribution of Type S error: $\text{Bernoulli}(\hat{\gamma}(\tilde{x}))$ (technically, its relative entropy). The maximization of the information gain is equivalent to minimizing the posterior entropy and consequently the log-loss (Settles, 2012).

3.4. Counterfactual Elicitation

Assuming it is possible to acquire (noisy) observations about counterfactuals in the training data, we can do more, as discussed in the introduction. Denote by $\mathbb{D}$ the set of training examples $\{x_i, a_i, y_i\}_{i=1}^n$ for which the factual outcomes y_i have been observed, and denote by $\mathbb{U}$ the counterfactual examples $\{x_i, 1 - a_i\}_{i=1}^n$, for which the outcomes are unknown. Then we can use active learning to construct a set $\mathbb{L}$ that contains the new observations, and $\mathbb{L}$ will be data which would not normally be available. At each iteration, the algorithm selects $\{x^*, a^*\} \in \mathbb{U}$ to solicit a counterfactual outcome y^*. After this, $\{x^*, a^*, y^*\}$ is added to $\mathbb{L}$, and removed from $\mathbb{U}$. So the optimization problem at each query iteration k becomes

$$x^*, a^* = \underset{\{x,a\} \in \mathbb{U}}{\arg\min}\, \mathbb{E}_{\hat{p}(y|x,a,\mathbb{D},\mathbb{L})}\left[\hat{\gamma}_{k+1}(\tilde{x})\right], \text{ where}$$

$$\hat{\gamma}_{k+1}(\tilde{x}) = \hat{p}\left(y[a_{\tilde{x}}] < y[1 - a_{\tilde{x}}] \mid \tilde{x}, \mathbb{D}, \mathbb{L}, \{x, a, y\}\right),$$

and $a_{\tilde{x}}$ is the treatment with the highest expected outcome for $\tilde{x}$: $a_{\tilde{x}} = \arg\max_{a'} \mathbb{E}_{\hat{p}(y[a']|\tilde{x},\mathbb{D},\mathbb{L},\{x,a,y\})}[y[a']]$.

4. Experiments and Results

We evaluate the performance of decision-making aware active learning (D-M Aware) using counterfactual elicitation in two cases. We compare the performance of D-M aware to two widely-used earlier active learning approaches, uncertainty sampling and maximum expected information gain (Culotta & McCallum, 2005; Roy & McCallum, 2001) (EIG), and also include results for a previously introduced special case of decision-making aware active learning, targeted-IG (Sundin et al., 2018). In order to make the methods comparable, we use the variant of D-M aware that also explores (see Section 3.3), as do the EIG and targeted-IG. The performance of the non-exploring variant of D-M aware is slightly lower (not shown) in all cases.

4.1. Simulated Data with Binary Outcomes

In this experiment, we study the proposed active learning approach in simulated data, where binary outcome y indicates the occurrence of an adverse event, and the decision-making task is therefore to choose the treatment that results in a lower probability of the adverse outcome. The setting is similar to that in the Fig. 1.

We model the probability of an adverse outcome given covariates $x \in \mathbb{R}$ and action $a \in \{0, 1\}$ using logistic regression with interactions and radial basis functions for local effects: $p(y \mid x, a) \sim \text{Bernoulli}(\theta_{x,a})$.The model is fit using a probabilistic programming language Stan (Stan Development Team, 2017; Carpenter et al., 2017).

Fig. 2(a) shows that the decision-making aware active learning criteria, D-M aware and Targeted-IG, are the fastest to improve decision-making performance, and achieve significant increase in correct decisions after just one query in two cases out of three. After five solicited counterfactuals, Expected Information Gain (EIG) has reached comparable but still lower performance.

4.2. Medical Semi-Synthetic Data with Continuous Outcomes

In the second experiment, the decision-making task is deciding on medical interventions in real medical data with continuous-valued synthetic outcomes.

We use the Infant Health and Development Program (IHDP) dataset from Hill (2011), also used e.g. by Shalit et al. (2017) and Alaa & van der Schaar (2017), including synthetic outcomes, containing 747 observations of 25 features. We evaluate the performance in leave-one-out cross-validation, but in order to make the problem even more realistically hard, for each of the 747 target units we choose randomly 100 observations as training examples.

We fit separate GPs to the outcomes of each treatment with GPy[1] (version 1.9.2), and use mixed noise likelihood to learn the noise in the observations acquired by active learning. We use an exponentiated quadratic kernel with a separate length-scale parameter for each variable, and optimize the hyperparameters using marginal likelihood. We use Gauss-Hermite quadrature of order 32 to approximate the expectations in D-M aware, Targeted-IG, and EIG.

The results in the IHDP data are similar as in the simulated data; The decision-making performance improves fastest with D-M aware and Targeted-IG, compared to EIG and uncertainty sampling (see Fig. 2(b)). D-M aware and Targeted-IG achieve statistically significant improvement in decision-making performance already with one query (based on 95% bootstrap confidence intervals).

4.3. Comparative Feedback

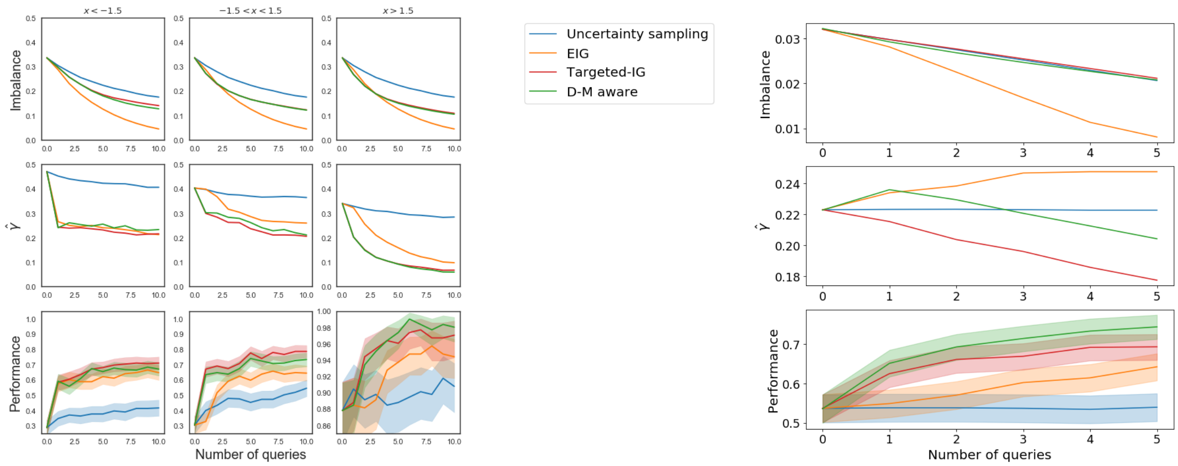
Last, we demonstrate the use of comparative observations for counterfactual elicitation, that is, indication of which potential outcome is higher. The motivation is that this type of feedback may be more natural for human experts to give. The setting is the same as in Section 4.1, with the difference that the query is about which treatment has lower Bernoulli parameter. We fit the outcomes with GPs using Stan (Stan Development Team, 2017; Carpenter et al., 2017), which allows the model to learn both from direct and comparative observations. The results in Fig. 3 show that comparative observations increase decision-making performance efficiently in the simulated data setting. We note that the results with comparative feedback are better than those in Fig. 2(a), because here the queries give information about $\theta_{x,1} > \theta_{x,0}$, thus providing more information than the direct observations on outcome $y[a] \mid x \sim \text{Bernoulli}(\theta_{x,a})$.

5. Discussion and Conclusions

As machine learning systems are being integrated into human decision-making workflows, it is increasingly important that deployed models are correct and reliable. In this paper we focused on the effect that imbalance in the training data can have on the reliability of comparisons of $\hat{p}(y[a] \mid \mathbf{x})$. We propose to improve the reliability by active learning that aims at maximizing the estimated reliability. In our experiments, this decision-making aware active learning outperforms standard methods in decreasing the error rate in decision-making.

One interesting application of the method is in personalized medicine, where counterfactual elicitation could solicit practitioners' knowledge to the model. For example, a clinician in a hospital has access to medical records of previous patients, and may also have personal experience about some

[1]Toolbox available at: https://sheffieldml.github.io/GPy



(a) (b)

Figure 2. Comparison of active learning criteria as a function of number of queries. (a) Simulated data (b) IHDP data. The topmost panel in each shows imbalance, middle panel the estimated Type S error rate, and the lowest panel the proportion of correct decisions. Information-gain-based approaches are more effective than uncertainty sampling, and the decision making-aware criteria D-M aware and Targeted-IG are the best. Shaded areas show the 95% bootstrap confidence interval.

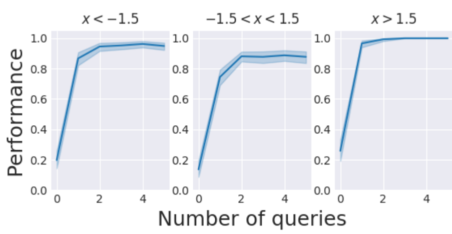


Figure 3. Comparative observations on the Bernoulli parameter in the simulated data is effective in increasing the proportion of correct decisions. The results are averaged over 3 target units in each response type regions and 100 repetitions.

of the previous patients. These data are rarely included in training sets of the models, but active learning and counterfactual elicitation could allow leveraging this additional source of information to infer more accurately about the future.

Acknowledgements

This work was supported by the Academy of Finland (grants 298742, 313122, 319264 and 294238), by Alfred Kordelin Foundation and by Jenny and Antti Wihuri Foundation. We acknowledge the computational resources provided by Aalto Science-IT project, and the support from the Finnish Center for Artificial Intelligence (FCAI).

References

Alaa, A. M. and van der Schaar, M. Bayesian inference of individualized treatment effects using multi-task Gaussian Processes. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems 30*, pp. 3424–3432. Curran Associates, Inc., 2017.

Bang, H. and Robins, J. M. Doubly robust estimation in missing data and causal inference models. *Biometrics*, 61 (4):962–973, 2005.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1):1–32, 2017. URL https://www.jstatsoft.org/v076/i01.

Culotta, A. and McCallum, A. Reducing labeling effort for structured prediction tasks. In *AAAI*, volume 5, pp. 746–751, 2005.

D'Amour, A., Ding, P., Feller, A., Lei, L., and Sekhon, J. Overlap in observational studies with high-dimensional covariates. *arXiv preprint arXiv:1711.02582*, 2018.

Funk, M. J., Westreich, D., Wiesen, C., Strmer, T., Brookhart, M. A., and Davidian, M. Doubly robust estimation of causal effects. *American Journal of Epidemiology*, 173(7):761–767, 2011.

Gelman, A. and Hill, J. *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press, 2007.

Gelman, A. and Tuerlinckx, F. Type S error rates for classical and Bayesian single and multiple comparison procedures. *Computational Statistics*, 15(3):373–390, 2000.

Hernán, M. and Robins, J. *Causal Inference*. Chapman & Hall, 2018.

Hill, J. L. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011.

Hirano, K., Imbens, G. W., and Ridder, G. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189, 2003.

Imbens, G. and Rubin, D. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

Javdani, S., Chen, Y., Karbasi, A., Krause, A., Bagnell, D., and Srinivasa, S. S. Near optimal Bayesian active learning for decision making. In *AISTATS*, volume 2, pp. 17, 2014.

Johansson, F., Shalit, U., and Sontag, D. Learning representations for counterfactual inference. In *International Conference on Machine Learning (ICML)*, 2016.

Lunceford, J. K. and Davidian, M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in medicine*, 23(19):2937–2960, 2004.

Mitchell, S., Potash, E., and Barocas, S. Prediction-based decisions and fairness: A catalogue of choices, assumptions, and definitions. *arXiv preprint arXiv:1811.07867*, 2018.

Morgan, S. and Winship, C. *Counterfactuals and causal inference*. Cambridge University Press, 2014.

Pearl, J. *Causality: models, reasoning and inference*. Cambridge University Press, 2009.

Rosenbaum, P. R. and Rubin, D. B. Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American statistical Association*, 79 (387):516–524, 1984.

Roy, N. and McCallum, A. Toward optimal active learning through Monte Carlo estimation of error reduction. In *International Conference on Machine Learning (ICML)*, pp. 441–448, 2001.

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

Saar-Tsechansky, M. and Provost, F. Decision-centric active learning of binary-outcome models. *Information Systems Research*, 18(1):4–22, 2007.

Schulam, P. and Saria, S. Reliable decision support using counterfactual models. In *Advances in Neural Information Processing Systems*, pp. 1697–1708, 2017.

Settles, B. Active learning: Synthesis lectures on artificial intelligence and machine learning. *Long Island, NY: Morgan & Clay Pool*, 2012.

Shalit, U., Johansson, F. D., and Sontag, D. Estimating individual treatment effect: generalization bounds and algorithms. In *International Conference on Machine Learning*, pp. 3076–3085, 2017.

Stan Development Team. Pystan: the python interface to stan, version 2.16.0.0, 2017. URL `http://mc-stan.org`.

Sundin, I., Peltola, T., Micallef, L., Afrabandpey, H., Soare, M., Mamun Majumder, M., Daee, P., He, C., Serim, B., Havulinna, A., Heckman, C., Jacucci, G., Marttinen, P., and Kaski, S. Improving genomics-based predictions for precision medicine through active elicitation of expert knowledge. *Bioinformatics*, 34(13):i395–i403, 2018.

Xu, Y., Xu, Y., and Saria, S. A Bayesian nonparametric approach for estimating individualized treatment-response curves. In *Machine Learning for Healthcare Conference (MLHC)*, pp. 282–300, 2016.

Yan, S., Chaudhuri, K., and Javidi, T. Active learning with logged data. In *International Conference on Machine Learning*, pp. 5517–5526, 2018.