Personalized Recommendation System Simulation Evaluation Technical Guidance Document

Haoda Fu's draft, please comments and contribute ***

December 14, 2018

Abstract

Key words and Phrases: *** ***.
Short title: ***

1 Introduction

Using data to generate actionable insights to guide patients, physicians, and payers to maximize their outcomes is a core topic for real world evidence research, and digital health solutions. Various solutions have been proposed, and it is crucial to understand their performance before applying it to make significant decisions for patients. This document provides guidance on how to design a simulation platform for such evaluation.

The evaluation is based on the potential outcome framework, i.e. the Rubin causal model Holland (1986). The following of this document is organized as below. In section 2, we provide an overview on how the data is generated. In section 3, we put data generation model in a unified framework. In section 4, we explain examples of training and testing data to explain in details on how to use them. In section 5, we provide details on how we evaluate the model performance. In the end, this document also provides some simulation settings as a reference.

2 Data Generate Process

Data for personalized recommendation system should contain 3 parts: contextual information ($X \in \mathcal{X}$), contextual decision/action ($A \in \mathcal{X}$), and reward ($Y \in \mathbb{R}$). The research question is how we can learn from such data to develop a rule $\mathcal{D}(\cdot)$ which is a map from $\mathcal{X}$ to $\mathcal{A}$. The ideal system recommends the optimal decisions for each patients, i.e.,

$$\mathcal{D}^*(X) = \underset{a \in \mathcal{A}}{\operatorname{argmax}}\, E\{Y|X, A = a\}.$$

To evaluate the model performance, each simulation generates training and testing data sets. The training data are used for model building. Users can further split training data for some tuning/validating purpose. The testing data sets are used to evaluate the model performance. The testing data sets contain the potential outcomes, and details will be provided in section 4.

There are three parts of data to generate in a sequence: X, A, and Y.

For contextual information X, they contains three types of variables: continuous variables (e.g. age, BMI, blood glucose values), ordinal categorical variables (e.g. cancer stage, severity of disease), or nominal categorical variables (e.g. gender, race) . The action A can be category variables, such as treatment, types of actions, or it can be continuous, such as dose. The reward/response Y can be continuous or ordinal categorical variable.

The action A can be depend on X, or it can be independent on X. For randomized control trials, treatment A is independent with X. For observational studies, treatment A often depends on X.

The reward/response Y is simulated based on different data generation models, and

details will be discussed in section 3. We use $Y(a)$ to denote potential outcome when a subject takes action a.

We have the following assumptions when we generate data,

- Stable unit treatment value assumption (SUTVA): $Y = \sum_{a=1}^{k} Y(a)I(A = a)$.
- No unmeasured confounder: $A \perp \{Y(a) : a \in \mathcal{A}\}|X$.

For the randomized control trials, the unmeasured confoundering assumptions are automatically satisfied.

3 Data Generation Models

Generalized Linear Model: When the responses are from exponential family, we generate data from the following GLM,

$$\ell\{E(Y|X)\} \quad = \quad \beta_0 + g(X) + t(A) + d(X, A),$$

where $\ell(\cdot)$ is a monotone link function. This model covers responses from various distributions including normal, binomial, Poisson, gamma, inverse gamma, and multinomial distributions.

Transformed Response Model: The transformed response model has the following form,

$$\tau(Y) \quad = \quad \beta_0 + g(X) + t(A) + d(X, A) + \epsilon,$$

where τ is a monotone transformation function, and $\epsilon \sim (0, \sigma^2)$ which can be nonparametric, i.e. mean equal to 0 with a finite second moment. This model covers the semi-parametric accelerated failure time model to generate time to event outcomes, and it also covers some often used PK/PD analytic models.

Intensity Function Model: When responses are time to event or recurrent events, the hazard or intensity function is often modeled as,

$$\lambda_i(t) \quad = \quad \lambda_0(t)\gamma_i \exp\{g(X_i) + t(A) + d(X_i, A)\},$$

where $\lambda_0(t)$ is a baseline hazard or intensity function, and γ_i is a frailty term with $\gamma_i \sim \text{Gamma}(1, \sigma^2)$. This model covers the Cox model, Anderson-Gail model, gamma-frailty model etc.

Remarks:

- $g(X)$ represents the prognostic effects. It can be parametric, e.g. $g(X) = X\beta$, semi-parametric, nonparametric additive model etc..

- $t(A)$ is the treatment effects.
- $d(X, A)$ models the predictive effects. Similarly, it can be parametric, semiparametric, or nonparametric models.

005_t03o1linear_Training_1.csv:

- 005: Dataset ID
- t03: 3 treatments
- o1: data from 1 observational study. (o: observational study, r:randomized control trial, m: mixture, e.g. m3 means that data from 3 studies, and some of them are observational studies and some of them are randomized control trials, study ID is considered as a covariate in X).
- linear: brief description of decision rule.

For each model, we have two data sets for testing and evaluation purposes. These two data sets are matched by patient ID. The first data set contains patients covariates, and it is labeled as "`nnn_xxxx_Testing_1.csv`". The second data set contains outcomes for evaluation which is labeled as "`nnn_xxxx_Testing_2.csv`".

[*Put Table 2 about here*]

[*Put Table 3 about here*]

4 Data Sets

We have 3 data sets, and they are training dataset M_1, testing data covariates M_2, testing data results M_3.

Example of M_1:

[*Put Table 1 about here*]

Example of M_2:

[*Put Table 2 about here*]

Example of M_3:

[*Put Table 3 about here*]

5 Evaluation Process

Suppose, we would like to evaluate a model which is a decision rule to map patient covariates into a treatment, i.e. $\mathcal{D}(\cdot) : X \mapsto A$.

Step 1: The input is M_1, and the output is a trained decision rule $\widehat{\mathcal{D}}(\cdot)$.

Step 2: The inputs are $\widehat{\mathcal{D}}(\cdot)$ and M_2, and the output is a vector $\widehat{A}$ with treatment assignment for each subject.

Step 3: The inputs are $\widehat{A}$ and M_3, and the outputs are the scores based on different evaluation criteria.

6 Evaluation Criteria

This subsection describes a high level evaluation process.

Primary criteria: The average benefit $N^{-1}\sum_{i=1}^{N}\sum_{a=1}^{k} Y_i(a) I\{a = \widehat{A}_i\}$.

Secondary criteria 1: Proportion of misclassification $N^{-1}\sum_{i=1}^{N} I\{A_i^o \neq \widehat{A}_i\}$.

Secondary criteria 2: Average of proportion of misclassification

$$k^{-1}\sum_{a=1}^{k}\left\{\sum_{i=1}^{N} I(A_i^o = a)\right\}^{-1}\sum_{i=1}^{N} I\{A_i^o \neq \widehat{A}_i, A_i^o = a\}.$$

Appendix A: ***

Naming format

`001_x4t2y1_dichotomous_linear1_train.csv`

- `001`: Dataset ID
- `x4`: The dimension of the covariate X is 4
- `t2`: 2 treatments
- `y1`: The dimension of Y is 1
- `dichotomous`: Description of the subgrouping criterion of X.
- `linear`: Description of the link function in the formulation of Y, e.g. `logit` for binary Y.

- `1`: Subcase ID
- `train`: Training set. `test_X` for testing data with only covariates; `test_Ys` for the response of the testing set, including the response of all the responses and the optimal decision Y^0

Simulation Settings and True Parameters (Confidential)

x4t2y1_dichotomous_linear1

This implements the simulation setting described in the "Improve numerical stability" part of Fu et al. (2016). The dimension of X is 4 and each covariate is generated from a uniform distribution $U(0, 1)$. The treatment A is a binary random variable coded as 1 and 2. It is randomly assigned with a probability as 0.5. The response Y is generated from the following model:

$$Y = \beta_0 + \text{sign}(X_2 - 0.5) + (A - 1) \cdot I_{X_1 \leq 0.6} + (2 - T) \cdot I_{X_1 > 0.6},$$

where β_0 is an arbitrary number, default as 5.

x4t2y1_dichotomous_linear2

This implements the simulation setting described in the "Convergence in RCTs" part of Fu et al. (2016). The dimension of X is 4 and each covariate is generated from a uniform distribution $U(0, 1)$. The treatment A is a binary random variable coded as 1 and 2. It is randomly assigned with a probability as 0.5. The response Y is generated from the following model:

$$Y = \text{sign}(X_2 - 0.5) + \theta \cdot (A - 1) \cdot I_{X \in D} + \theta \cdot (2 - A) \cdot I_{X \in D^C} + \epsilon,$$

where ϵ is i.i.d. $N(0, 1)$. The true subgroup D is defined by different depths from 1 to 3 as

- When depth is equal to 1, $A = X : X_1 <= 0.6$
- When depth is equal to 2, $A = X : X_1 <= 0.7 \cap X_3 > 0.3$
- When depth is equal to 3, $A = X : X_1 <= 0.8 \cap X_2 > 0.2 \cap X_3 > 0.2$.

θ is used to change the signal-noise ratio, ranging from 0.1 to 0.5, default as 0.3.

x5t2y1_dichotomous_linear1

This mplements the simulation setting described in the "Convergence in observational studies" part of Fu et al. (2016). The dimension of X is 5 and each covariate is generated from a uniform distribution $U(0,1)$. The treatment A is a binary random variable coded as 1 and 2. It is randomly assigned with a probability from the following model:

$$\text{logit}(p) = -0.5b + bX_i,$$

where b and i are arbitrary, default as 6.5 and 1. The response Y is generated from the following model:

$$Y = 1 + 2 \cdot X_2 + \theta \cdot (A-1) \cdot I_{X_1>0.5} + \theta \cdot (2-A) \cdot I_{X_1 \leq 0.5} + \epsilon,$$

where ϵ is i.i.d. $N(0,1)$, θ is 0.5 by default.

x10t2y1_dichotomous_linear

This implements the simulation settings described in the "Variable importance" part of Fu et al. (2016). The dimension of X is 10 and each covariate is generated from a uniform distribution $U(0,1)$. The treatment A is a binary random variable coded as 1 and 2. It is randomly assigned with a probability of 0.5. The response Y is generated from the following model:

$$Y = \text{sign}(X_2 - 0.5) + 0.5(A-1) \cdot I_{X \in D} + 0.5 \cdot (2-A) \cdot I_{X \in D^C} + \epsilon,$$

where ϵ is i.i.d. $N(0,1)$, $A = X : X_4 \leq 0.7 \cap X_6 \leq 0.7$.

x3t2y1_torus_linear

$X \in \mathbb{R}^3$, and each covariate is generated from a uniform distribution $U(-1,1)$. The treatment A is a binary random variable coded as 1 and 2. It is randomly assigned with a probability of 0.5. The response Y is generated from the following model:

$$Y = \beta + 3 \cdot X_3 + 0.3 \cdot (A-1) + 0.2 \cdot (2-A) + \theta \cdot (A-1) \cdot I_{X \in D} + \theta \cdot (2-A) \cdot I_{X \in D^C} + \epsilon,$$

where ϵ is i.i.d. $N(0,1)$. β, θ are 1 and 0.8 by default. D is a 3-D torus described by:

$$(\sqrt{X_1^2 + X_2^2} - R)^2 + X_3^2 \leq r^2,$$

where R and r are the major radius and minor radius of the torus respectively. The shape of the torus is illustrated below:

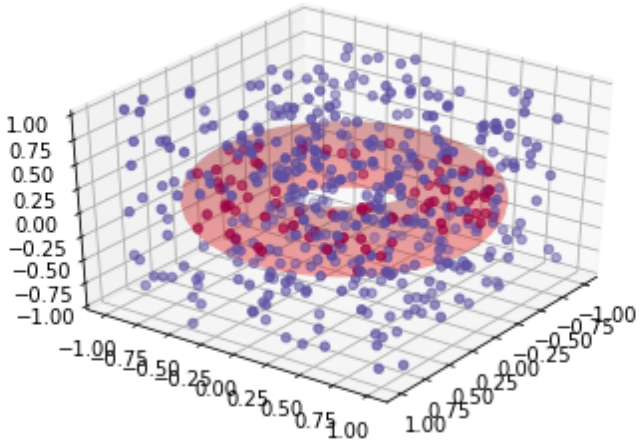

Figure 1: Illustration of a torus in $\mathbb{R}^3$. The blue and red dots denote the two true subgroups of X

x11t2y1_complex_logit

This implements the simulation settings described in Xu et al. (2015). The data contains 5 binary covariates X_1 toX_5, 5 ordinal covariates X_6 to X_{10} and 1 continuous covariateX_{11}. The binary variables are generated by Bernoulli distribution with successful probability of 1/2. All ordinal covariates have 4 uniformly distribution levels. The continuous covariate is simulated form the standard normal distribution. The treatment A is independent of the covariates and takes value of 1 and 2 with equal probability. The binary response Y is generated using logit as link function. The last letter in the file name corresponds to the case label in the original paper which indicates different model to generate the response Y. The detailed formulas for logit$P(Y = 1)$ are described in the following:

A. $0.5 \cdot I_{X_3=1} + 2(I_{X_2=1} + I_{X_6=3} \cdot I_{X_1=1}) \cdot (2A - 3)$;

D. $\log\log(I_{X_7=3} + I_{X_8=3} + 5(I_{X_6=2} + I_{X_1=1,X_2=1}) \cdot (2A - 3) + 20)^2$;

F. $0.5I_{X_1=1} + 0.5I_{X_2=1} + 2I_{X_{11}<5,X_6<2} \cdot (2A - 3)$.

The term $(2A - 3)$ transforms the label of A from 1/2 to -1/1.

x4t2y1_discrete_linear

This implements the simulation settings described in Su et al. (2009). Each set contains 4 covariates X_1 to X_4 simulated from a discrete uniform distribution over (0.02, 0.04, ..., 1.00). The treatment A is independent of the covariates and takes value of 1 and 2 with

equal probability. The continuous response Y is generated based on different models which are described below:

B. $Y = 2 + 2 \cdot (A-1) + 2Z_1 + 2Z_2 + 2 \cdot (A-1) \cdot Z_1 Z_2 + \epsilon, \quad \epsilon \sim N(0,1)$;

D. $Y = 10 + 10 \cdot (A-1) \cdot \exp\{(X_1 - 0.5)^2 + (X_2 - 0.5)^2\} + \epsilon, \quad \epsilon \sim N(0,1)$;

E. $Y = 2+2\cdot(A-1)+2Z_1+2Z_2+2\cdot(A-1)\cdot Z_1+2\cdot(A-1)\cdot Z_2+\epsilon, \quad \epsilon \sim Unif(-\sqrt{3}, \sqrt{3})$;

F. $Y = 2 + 2 \cdot (A-1) + 2Z_1 + 2Z_2 + 2 \cdot (A-1) \cdot Z_1 + 2 \cdot (A-1) \cdot Z_2 + \epsilon, \quad \epsilon \sim exp(1)$.

The term $A-1$ transforms the label of A from 1/2 to 0/1 and $Z_1 = I_{X_1 \leq 0.5},\ Z_2 = I_{X_2 \leq 0.5}$

x10t2y1_plane_linear

Implement the simulation setting 1 described in Wang et al. (2018). The dimension of X is 10 and each covariate is generated from a uniform distribution $U(0,1)$. The treatment A is a binary random variable coded as -1 and 1. It is randomly assigned with a probability of 0.5. The response Y is generated from the following model:

$$Y = 1 - 2X_1 + X_2 - X_3 + 2(1 - X_1 - X_2)A + \epsilon_Y,$$

where ϵ is i.i.d. $N(0,1)$.

x10t2y1_circle_linear

Implement the simulation setting 2 described in Wang et al. (2018). The dimension of X is 10 and each covariate is generated from a uniform distribution $U(0,1)$. The treatment A is a binary random variable coded as -1 and 1. It is randomly assigned with a probability of 0.5. The response Y is generated from the following model:

$$Y = 1 - 2X_1 + X_2 - X_3 + 8(1 - X_1^2 - X_2^2)A + \epsilon_Y,$$

where ϵ is i.i.d. $N(0,1)$.

x4t2y1_checkboard_linear

Implement the motivation example about checkboard described in Zhu et al. (2015). The dimension of X is 4 and each covariate is generated from a uniform distribution $U(0,1)$. The treatment A is a binary random variable coded as -1 and 1. It is randomly assigned with a probability of 0.5. The response Y is generated from the following model:

$$\begin{aligned} Y &= A \times (I_{X \in D_1} - I_{X \in D_1}), \\ D_1 &= \big\{X | \{X_1 < 1/3\} \cap \{X_1 > 2/3\}\big\}, \\ D_2 &= \big\{X | \{X_2 < 1/3\} \cap \{X_2 > 2/3\}\big\}. \end{aligned}$$

x2t2y1_spiral_linear

Implement the motivating subgroup structure of spiral shape. The dimension of X is 2 and each covariate is generated from a uniform distribution $U(0,1)$. The treatment A is a binary random variable coded as -1 and 1. It is randomly assigned with a probability of 0.5. Denote $tan(\theta) = \frac{X_2}{X_1}$ and $r = \sqrt{X_1^2 + X_2^2}$. The response Y is generated from the following model:

$$
\begin{aligned}
&Y = A \times (I_{X \in D_1} - I_{X \in D_1}),\\
&D_1 = \big\{X | 0.1 \leq r - 1.5(\theta + (2k + 0.5)\pi) \leq 2,\ k = 0, 1, 2\big\},\\
&D_2 = \big\{X | 0.1 \leq r - 1.5(\theta + (2k + 1.5)\pi) \leq 2,\ k = 0, 1, 2\big\}.
\end{aligned}
$$

The spiral shape is illustrated below:

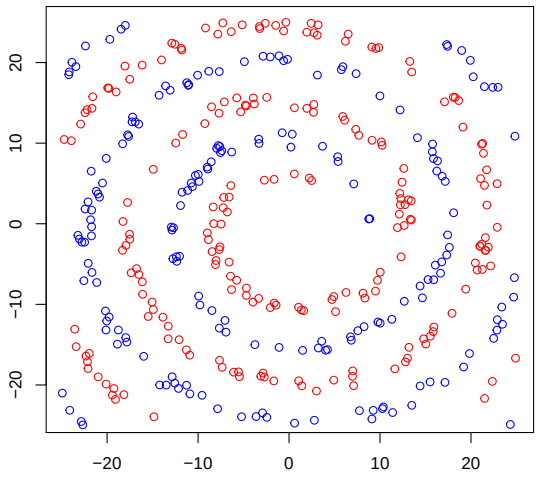

Figure 2: Illustration of the spiral subgroup. The blue color and red color represent two subgroups

x10t2y1_cox_circle

Implement the Cox model with circle subgroup structure. The dimension of X is 10 and each covariate is generated from a uniform distribution $U(0,1)$. The dimension of truly associated covariates is 4. The treatment A is a binary random variable coded as -1 and 1. It is randomly assigned with a probability of 0.5. The survival time T is generated as following: we set the constant baseline hazard as 1.

$$T \sim \exp(1 - 2X_1 + X_2 - X_3 + 8(1 - X_1^2 - X_2^2)A)$$

x1t8y3_Q_survival

Implement a simplified version of the simulation setting described in Goldberg and Kosorok (2012). We use the original notation for convenience. Denote the discrete time point $i = 0, 1, 2, 3$, the tumor size at time i as $T(i) \in [0, 1]$ and the wellness at time i as $W(i) \in [0.25, 1]$. The trial begin with $T(0) = 1$ and $W(0)$ uniformly distributed on [0.5,1]. At each time point i, we consider two optional treatments A and B. The immediate effects of A are:

$$W(i^+|A) = W(i) - 0.5, \tag{A.1}$$
$$T(i^+|A) = T(i)/8W(i)), \tag{A.2}$$

where $W(i^+|A)$ is the wellness at i after treatment A. Similarly the immediate effects of the treatment B are:

$$W(i^+|B) = W(i) - 0.25, \tag{A.3}$$
$$T(i^+|B) = T(i)/(4W(i)). \tag{A.4}$$

The wellness and the tumor size of the next time point are computed as:

$$W(i+1) = \max(\ W(i^+) + (1 - W(i^+))(1 - 2^{-1/2}) + \epsilon, 0.25\ ), \tag{A.5}$$
$$T(i+1) = \min(\ T(i^+) + 4T(i^+)/3 + \epsilon, 1\ ), \tag{A.6}$$

where $\epsilon \sim N(0, 0.1)$ For $i = 0, 1, 2$, two treatments are assigned to the patient with equal probability independently. We use $A_1A_2A_3$ to denotes the treatment sequence where $A_i \in A, B$. We encode all the eight probable treatments into $1, 2, ..8$, i.e., $1 = AAA$, $2 = AAB$, $3 = ABA$, $4 = BAA$, $5 = ABB$, $6 = BBA$, $7 = BAB$, $8 = BBB$. So the action is eight dimensional. Besides, we define one dimensional $X = W(0)$. We also assume the survival function of the patient as an exponential distribution with mean $S(i) = 3(W(i^+)+2)/(20T(i^+))$, $i = 0, 1, 2$. We define three dimensional $Y = [S(0)\ S(1)\ S(2)]$ and the optimal action is chosen to maximize $\sum_{i=0}^{2} S(i)$.

x11t3y10_longitudinal_log

This part implement the simulation setting described in Zhu and Qu (2016). For each individual, it follows a longitudinal log-linear mixed-effect model:

$$\log Y_j = X_j\beta + Z_j b + d\log(D_j) + \epsilon_j, i = 1, \cdots, N,\ j = 1, \cdots, 10, \tag{A.7}$$

where Y_j is a continuous post-treatment response of interest at time j; A_j is the drug dosage at time point j; and d is the corresponding fixed coefficient. The time-varying fixed-effect covariates X_j are generated from Unif$(0.75, 2.25)$; the fixed-effect parameters $\beta = 0.5$ and $d = 1.2$; The covariate Z_j is time-invariant ($Z_1 = \cdots = Z_{ij}$) with $Z_1 \sim U(0.75, 2.25)$. For

the generation of the random-effect b, half of the subjects are from Unif$(-1.5, -0.5)$, and another half from Unif$(1.5, 0.5)$. For convenience, The dosage assigned is sampled from three different level (100,300,600) and assumed to be time-invariant: $D_1 = \cdots = D_j$. The error term $\boldsymbol{\epsilon_i}$ follows $N(\mathbf{0}, 0.5\boldsymbol{R})$, where R is AR(1) correlation matrix with the correlation coefficient $\rho = 0.8$.

To keep consistent with the notation used before, we define $\boldsymbol{X} = (X_1, \cdots, X_{10}, Z_1) \in \mathbb{R}^{11}$ and $\boldsymbol{Y} = (Y_1, \cdots, Y_1 0) \in \mathbb{R}^{10}$ for each individual. And the dosage level (100,300,600) is coded as $A = 1, 2, 3$ respectivley. The optimal dosage A^0 is chosen to maximize $\sum_{j=1}^{10} Y_j$.

x10t3y1_linear_plane

This setting extend the case in **x10t2y1_plane_linear** with binary treatments to the three classes of treatments. The dimension of X is 10 and each covariate is generated from a uniform distribution $U(0, 1)$. The dimension of true covariates is 3. The treatment A is a random variable uniformly sampled from $\{1, 2, 3\}$ representing three different treatments. The response Y is generated from the following model:

$$\begin{aligned} Y &= 1 - 2X_1 + X_2 - X_3 \\ &\quad + I_{\{X1+X2+X3<1.5\}} I_{\{A=1\}} + I_{\{1.5\leq X1+X2+X3<2.5\}} I_{\{A=2\}} + I_{\{X1+X2+X3\geq 2.5\}} I_{\{A=3\}} + \epsilon_Y, \end{aligned}$$

where ϵ_Y is i.i.d. $N(0, 1)$.

x10t3y1_linear_circle

This setting extend the case in **x10t2y1_circle_linear** with binary treatments to the three classes of treatments. The dimension of X is 10 and each covariate is generated from a uniform distribution $U(0, 1)$. The dimension of true covariates is 3. The treatment A is a random variable uniformly sampled from $\{1, 2, 3\}$ representing three different treatments. The response Y is generated from the following model:

$$\begin{aligned} Y &= 1 - 2X_1 + X_2 - X_3 \\ &\quad + I_{\{X1^2+X2^2+X3^2<1.5\}} I_{\{A=1\}} + I_{\{1.5\leq X1^2+X2^2+X3^2<2.5\}} I_{\{A=2\}} + I_{\{X1^2+X2^2+X3^2\geq 2.5\}} I_{\{A=3\}} + \epsilon_Y, \end{aligned}$$

where ϵ_Y is i.i.d. $N(0, 1)$.

References

Fu, H., J. Zhou, and D. E. Faries (2016). Estimating optimal treatment regimes via subgroup identification in randomized control trials and observational studies. *Statistics in medicine 35*(19), 3285–3302.

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

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

Su, X., C.-L. Tsai, H. Wang, D. M. Nickerson, and B. Li (2009). Subgroup analysis via recursive partitioning. *Journal of Machine Learning Research 10*(Feb), 141–158.

Wang, Y., H. Fu, and D. Zeng (2018). 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.

Xu, Y., M. Yu, Y.-Q. Zhao, Q. Li, S. Wang, and J. Shao (2015). Regularized outcome weighted subgroup identification for differential treatment effects. *Biometrics 71*(3), 645–653.

Zhu, R., D. Zeng, and M. R. Kosorok (2015). Reinforcement learning trees. *Journal of the American Statistical Association 110*(512), 1770–1784.

Zhu, X. and A. Qu (2016). Individualizing drug dosage with longitudinal data. *Statistics in medicine 35*(24), 4474–4488.

Table 1: *An illustration training dataset*

ID	Y	A	X_1	X_2	X_3	$\cdots$
1	1.5	1	0	26	7.8	$\cdots$
2	1.2	2	1	28	8.2	$\cdots$
3	0.3	3	1	31	8.9	$\cdots$
4	0.9	2	0	35	9.4	$\cdots$
5	1.7	1	1	22	7.3	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$
n	1.6	2	0	29	8.1	$\cdots$

Table 2: *An illustration testing dataset: nnn_xxxx_Testing_1.csv.*

ID	X_1	X_2	X_3	$\cdots$
1	0	26	7.8	$\cdots$
2	1	28	8.2	$\cdots$
3	1	31	8.9	$\cdots$
4	0	35	9.4	$\cdots$
5	1	22	7.3	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$
N	0	29	8.1	$\cdots$

Table 3: *An illustration testing dataset: nnn_xxxx_Testing_2.csv.* $Y(a)$ is the potential outcome taking treatment a, A is the observed treatment assignment, and A^o is the oracle optimal treatment assignment based on $d(X, A)$.

ID	$Y(1)$	$Y(2)$	$\cdots$	$Y(k)$	A	A^o
1	1.2	1.5	$\cdots$	1.3	3	2
2	1.3	1.1	$\cdots$	1.4	2	3
3	0.9	0.8	$\cdots$	1.7	1	3
4	1.8	1.6	$\cdots$	1.2	1	1
5	1.4	1.4	$\cdots$	1.5	2	2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
N	1.7	1.4	$\cdots$	1.1	3	1

Table 4: *An illustration testing dataset: nnn_xxxx_Testing_2.csv.* The observed benefit is $V = 1/N \sum(1.2 + 1.1 + 0.9 + 1.6 + 1.4 + \cdots + 1.4)$, and theoretical optimal value is $V^o = 1/N \sum(1.5 + 1.3 + 0.8 + 1.8 + 1.4 + \cdots + 1.7)$, and the estimated value is $\widehat{V} = 1/N \sum(1.5 + 1.3 + 0.9 + 1.6 + 1.4 + \cdots + 1.7)$.

ID	$Y(1)$	$Y(2)$	A	A^o	$\widehat{A}$
1	1.2	1.5	1	2	2
2	1.3	1.1	2	1	1
3	0.9	0.8	1	2	1
4	1.8	1.6	2	1	2
5	1.4	1.4	2	2	2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
N	1.7	1.4	2	1	1