EVALUATION OF ACTIVE FEATURE ACQUISITION METHODS UNDER MISSING DATA

Anonymous authors
Paper under double-blind review

ABSTRACT

Machine learning (ML) methods generally assume the full set of features are available at no cost. If the acquisition of a certain feature is costly at run-time, one might want to balance the acquisition cost and the predictive value of the feature for the ML task. The task of training an AI agent to decide which features are necessary to be acquired is called active feature acquisition (AFA). Current AFA methods, however, are challenged when the AFA agent has to be trained/tested with datasets that contain missing data. We formulate, for the first time, the problem of active feature acquisition performance evaluation (AFAPE) under missing data, i.e. the problem of adjusting for the inevitable missingness distribution shift between train/test time and run-time. We first propose a new causal graph, the AFA graph, that characterizes the AFAPE problem as an intervention on the reinforcement learning environment used to train AFA agents. Here, we discuss that for handling missing data in AFAPE, the conventional approaches (off-policy reinforcement learning, blocked feature acquisitions, imputation and inverse probability weighting (IPW)) often lead to biased results or are data inefficient. We then propose active feature acquisition importance sampling (AFAIS), a novel estimator that is more data efficient than IPW. We demonstrate the detrimental conclusions to which biased estimators can lead as well as the high data efficiency of AFAIS in multiple experiments using simulated and real-world data under induced MCAR, MAR and MNAR missingness.

1 INTRODUCTION

Machine learning methods generally assume the full set of input features are available at run-time at little to no cost. This is, however, not always the case as acquiring features may impose a significant cost. For example in medical diagnosis, the cost of feature acquisition (e.g. a biopsy test) could include both its monetary cost as well as the potential adverse harm for patients. In this case, the predictive value of a feature should be balanced against its acquisition cost. Physicians acquire certain features via biopsies, MRI scans, or lab tests, only if their diagnostic value outweighs their cost or risk. This challenge becomes more critical when physicians aim to predict a large number of diverse outcomes, each of which has different sets of informative features. Going back to the medical example, a typical emergency department (ED) is able to diagnose thousands of different diseases based on a large set of possible observations. For every new emergency patient entering ED with an unknown diagnosis, clinicians must narrow down their search for a proper diagnosis via step by step feature acquisitions. In this case a ML model designed to do prediction given the entire feature set is infeasible.

Active feature acquisition (AFA) addresses this problem by designing two AI systems: i) a so-called *AFA agent*, deciding which features must be observed, while balancing information gain vs. feature cost; ii) an ML prediction model, often a classifier, that solves the prediction task based on the acquired set of features. An AFA agent, by definition, induces missingness by selecting only a subset of features. We call this *AFA missingness* which occurs at run-time (e.g. when the AFA agent is deployed at the hospital). In addition, in many AFA applications, retrospective data which we use for model training and evaluation also contain missing entries. This is induced by a different feature acquisition process (e.g. by physicians, ordering from a wide range of diagnostic tests). We call this *observational missingness*. While using retrospective data (during training/evaluation), the agent can only decide among available features. At run-time, however, we make the assumption that

the agent has the freedom to choose from all features. This corresponds to a feature "availability" distribution shift that requires adjustment. Apart from difficulties of training an AFA agent using incomplete data, estimating the real world performance of agents (at run-time) using incomplete retrospective data is nontrivial and challenging. Evaluation biases might lead to false promises about the agent's performance, a serious risk especially in safety-critical applications.

This paper is the first, to our knowledge, to systematically study the evaluation of AFA agents under this distribution shift. We call this problem active feature acquisition performance evaluation (AFAPE) under missing data. Our work on the AFAPE problem contains 4 major contributions for the field of AFA:

First, we propose the *AFA graph*, a new causal graph, that characterizes the AFA problem under observational missingness as off-environment reinforcement learning (RL) under an environment intervention. This is, in our opinion, a very important connection between the fields of missing data and reinforcement learning that will result in cross-fertilization of ideas from these two areas.

Second, we show that the AFA literature only contains approaches to handle missing data that are either derived from a pure RL perspective (off-policy policy evaluation (OPE) (Chang et al., 2019), blocking of acquisition actions (Janisch et al., 2020; Yoon et al., 2018)) or correspond to biased methods for missing data (conditional mean imputation (An et al., 2022; Erion et al., 2021; Janisch et al., 2020)). We show that both approaches will almost certainly lead to biased results and/or can be extremely data inefficient. We demonstrate in experiments with missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR) patterns that these biased evaluation methods can lead to detrimental conclusions about which AFA agent performs best. This can lead for example to high risks for patients' lives if these methods are deployed without proper evaluation.

Third, we bring the readers' attentions to unbiased estimators from the missing data literature (inverse probability weighting (IPW) (Seaman & White, 2013) and multiple imputation (MI) (Sterne et al., 2009)) which haven't been applied to AFA previously. These methods, however, do not account for the special RL structure of the AFAPE problem which, as we show, leads to data inefficiency.

Fourth, we instead propose AFAIS (active feature acquisition importance sampling), a new estimator based on an off-environment bandit view. AFAIS is more data efficient than IPW, but cannot always be used for complex MNAR scenarios. For these cases, we propose a modification to AFAIS that allows it to be closer to IPW when required, while still being data efficient. We demonstrate the improved data efficiency of AFAIS over IPW in multiple experiments.

2 Related methods

AFA: Various approaches have been proposed for designing AFA agents and prediction models for active feature acquisition (AFA) (An et al., 2006; Li & Oliva, 2021a; Li et al., 2021; Chang et al., 2019; Shim et al., 2018; Yin et al., 2020). This work focuses, however, not on any particular AFA method, but on the evaluation of any AFA method under missingness. Nevertheless, we refer the interested reader to Appendix A.1 for a more detailed literature review of existing AFA methods and a distinction between AFA and other related fields.

Missing data: AFAPE under missingness can be viewed as a missing data problem, and hence methods from the missing data literature can be adopted. There are in general two difficulties for solving missing data problems. The first is identification, i.e. the determination whether the estimand of the full (unknown) data distribution can be estimated from the observed data distribution. The second is estimation, for which there exist generally two strategies which are based on importance sampling (IS), i.e. inverse probability weighting (IPW) (Seaman & White, 2013), and multiple imputation (MI) (Sterne et al., 2009). See Appendix A.2 for an in-depth review on missing data.

Off-policy policy evaluation (OPE): As we show in Section 3, the AFAPE problem can be formulated as an off-policy policy evaluation (OPE)(Dudik et al., 2011; Kallus & Uehara, 2020) problem. The goal in OPE is to evaluate the performance of a "target" policy (here the AFA policy) from data collected under a "behavior" policy (here the observational missingness induced by e.g. the doctor). Similar to missing data, there are two general OPE estimation strategies based on IS (Precup et al.,

2000) and outcome regression (using q-functions (Munos & Szepesvári, 2008)). As we will discuss in Section 3 these approaches are extremely data inefficient and can lead to biased results in complex missing data scenarios.

3 METHODS

We open this section by introducing a new view on missingness, the *sequence view*. It complements the traditional view on missingness which we denote as the *set view*. Using the sequence view, we formulate AFA as a sequential decision problem, and the observational missingness in AFA by a novel graphical representation, the *AFA graph*. We continue by formulating the AFAPE problem as a distribution shift problem. As a conclusion, this formulation allows us to first, discuss different views on the problem, second, study existing solutions, and finally, propose the AFAIS estimator, a novel estimator for AFAPE.

3.1 THE SET AND SEQUENCE VIEWS ON MISSINGNESS

The traditional view on missingness defines the problem by features X, missingness indicators $\bar{R}$ and observed proxies $\bar{X}$. In AFAPE, we also consider corresponding feature acquisition costs $\bar{C}$. The relationship between variables in this view is graphically modelled by a *missing data graph* (m-graph) (Figure 1A) (Mohan et al., 2013; Shpitser et al., 2015). We name this view the *set view*, as it focuses on the set of acquired features collectively. See Appendix A.2 for an in-depth review on this view on missingness.

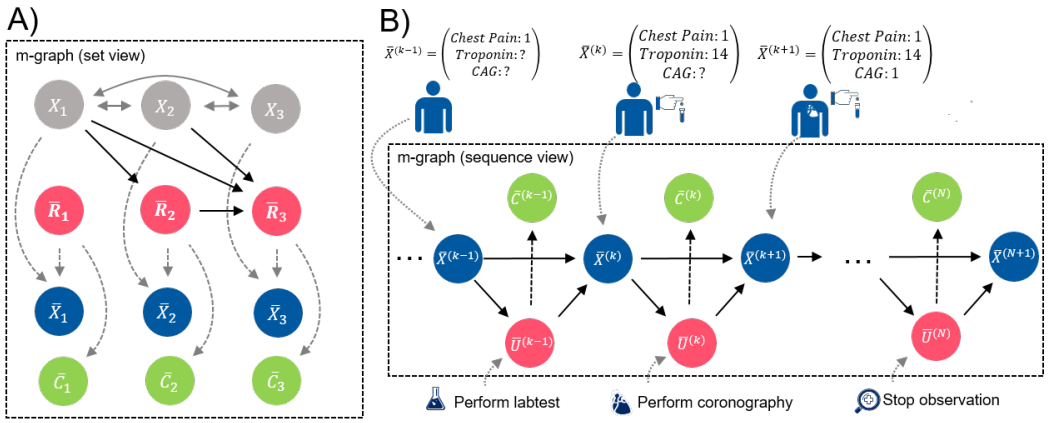

Figure 1: Missing data graphs for a simplified hospital setting of a heart attack diagnosis: X_1: chest pain; X_2: troponin (lab test); X_3: coronography (CAG) (invasive imaging). A) m-graph from a set view. Presence of chest pain increases the chance for the ordering of a lab test ($X_1 \rightarrow \bar{R}_2$). Based on the results from the lab test or the symptom of chest pain, a coronography might be ordered ($X_1 \rightarrow \bar{R}_3$; $X_2 \rightarrow \bar{R}_3$; $\bar{R}_2 \rightarrow \bar{R}_3$). The application of coronography produces a large cost $\bar{C}_3$ due to its invasiveness and radiation. B) m-graph from a sequence view, showing a (different) MAR missingness RL policy $\pi(\bar{U}^{(k)}|\bar{X}^{(k)}, \bar{R}^{(k)}) \in \mathcal{F}_{MAR}$. Based on the acquired features at acquisition step k, a new observation action $\bar{U}^{(k)}$ is performed (here: $\bar{U}^{(k)} =$ "observe(X_3)").

In order to formulate the AFA problem, we introduce a new view on missingness, which we denote as the *sequence view* due to its focus on the sequence of feature acquisitions (Figure 1B). The sequence view arises from inducing a topological ordering in m-graphs, as follows. We assume that a missingness indicator $\bar{R}$ is formed during a sequence of observation actions $\bar{U} = (\bar{U}^{(1)}, ..., \bar{U}^{(N)})$, where action $\bar{U}^{(k)} =$ "observe(X_i)" sets $\bar{R}_i = 1$ if $k < N$. The last observation action is always defined as $\bar{U}^{(N)} =$ "stop observation". Thus, $\bar{R}^{(k)}$ (resp. $\bar{X}^{(k)}$) corresponds to the intermediate missingness indicator (resp. intermediate observed proxy) after k steps of observation. We denote the distribution $p(\bar{U}^{(k)}|X, \bar{R}^{(k)}) \equiv \pi(\bar{U}^{(k)}|X, \bar{R}^{(k)})$ as the missingness RL policy. In order to introduce an AFA policy later, we need to define a restriction to the space of policies as follows.

We define the functional space $\mathcal{F}_{MAR}$ by restricting the missingness RL policy to be dependent only on data observed up to the current time step k: $\pi_{MAR}(\bar{U}^{(k)}|X, \bar{R}^{(k)}) = \pi_{MAR}(\bar{U}^{(k)}|\bar{X}^{(k)}, \bar{R}^{(k)})$. A MAR policy is denoted as $\pi_{MAR} \in \mathcal{F}_{MAR}$. Note that this is a special case of MAR in which observations depend on the earlier variables in the m-graph under the sequence view.

3.2 Problem Formulation: Active feature acquisition (AFA)

The AFA problem without observational missingness is visualized as a causal graph in Figure 2. The goal of AFA is to find i) an optimal AFA agent that decides which features to acquire and ii) a prediction model which uses the acquired features to perform a prediction. We assume here a classification task as the subsequent prediction task, but the problem extends naturally to other tasks.

Data (Figure 2 bottom right):

Consider a $n_D \times (d_x + 1)$ dataset comprised of n_D i.i.d. realizations of a d_x-dimensional random vector of features $X \in \mathbb{R}^{d_x}$ and a 1-dimensional random variable of the label $Y \in \mathbb{N}$. We consider X and Y fully observed for now and denote them as the environment of the AFA agent. The AFA graph shows an arrow $X \rightarrow Y$ (the features cause the label), but all results in this paper also hold under a reversed causal relationship ($X \leftarrow Y$).

AFA missingness (Figure 2 top left):
The AFA missingness is induced by the AFA agent that decides which subset of features to acquire. Let the missingness indicator $\hat{R}$ denote the AFA missingness in X (which produces observed proxies $\hat{X}$). The corresponding policy $\pi(\hat{U}^{(k)}|X^{(k)}, \hat{R}^{(k)})$ is the AFA missingness RL policy. Our MAR assumption holds for the AFA missingness RL policy (i.e. $\pi(\hat{U}^{(k)}|X^{(k)}, \hat{R}^{(k)}) = \pi(\hat{U}^{(k)}|\hat{X}^{(k)}, \hat{R}^{(k)}) \in \mathcal{F}_{MAR}$). Further, consider a predefined, fixed acquisition cost $c_{acq}(i)$ of observing a feature X_i. The *acquisition cost* random variable $\hat{C}$ represents the cost that the AFA policy exerts when observing features in X: $\hat{C} = \sum_{i=1}^{d_x} \hat{R}_i \cdot c_{acq}(i)$. We make one important assumption about the application of the AFA agent at run-time/deployment:

Assumption 1 *The AFA agent has full control of missingness at run-time.*

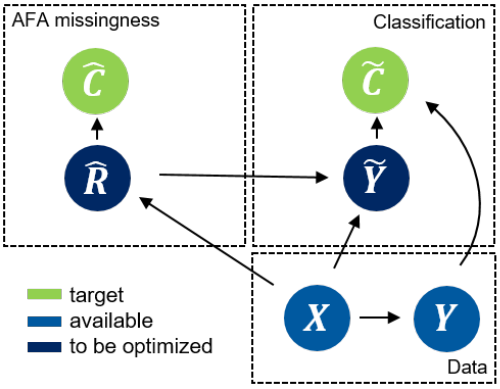

Figure 2: Causal graph for AFA: The AFA agent decides which of the features in X to acquire (which produces $\hat{R}$) and pays feature acquisition costs $\hat{C}$. Based on the acquired set of features, the classifier predicts $\tilde{Y}$ and pays a misclassification cost $\tilde{C}$ if Y and $\tilde{Y}$ mismatch.

Assumption 1 implies that there are no other mechanisms that prevent the feature acquisition when the AFA agent requests the acquisition of a feature, i.e. the features are available for the agent, when asked for.

Classification (Figure 2 top right):
Consider a predicted label random variable $\tilde{Y} = f(\hat{R}, X) = f(\hat{R}, \hat{X})$ that is predicted based on the observed features $\hat{X}$ using classifier f. Further, consider wrong predictions produce a misclassification cost $\tilde{C} = c_{mc}(\tilde{Y}, Y)$, where $c_{mc}(y_l, y_m) : \mathbb{N}^2 \rightarrow \mathbb{R}$ is the cost of predicting a label of class y_l as belonging to class y_m.

AFA objective:
The goal of AFA is to find a missingness RL policy $\pi_\theta(\hat{U}^{(k)}|\hat{X}^{(k)}, \hat{R}^{(k)})$, parameterized by θ and a classifier $f_\phi(\hat{R}, \hat{X})$, parameterized by ϕ, that minimize the total sum of acquisition and misclassification costs $C = \hat{C} + \tilde{C}$. The associated objective function is defined as

$$\arg\min_{\phi,\theta} J(\phi, \theta) = \arg\min_{\phi,\theta} \mathbb{E}\left[C \middle| \phi, \theta\right] \tag{1}$$

where $J(\phi, \theta)$ denotes the expected total cost after applying AFA missingness RL policy and classifier. Based on the described sequential nature of the AFA problem, one may find an optimal AFA agent using a method of choice, such as RL. In this work, however, we focus on the important issue of AFAPE, i.e. the unbiased evaluation of AFA agents (via estimating J) under observational missingness, which arises when the available retrospective dataset contains missingness. For ease of reference, we drop ϕ and θ, the arguments of the J notation.

3.3 Problem Formulation: Active Feature Acquisition Performance Evaluation (AFAPE) under Missing Data

The objective of evaluating AFA agents is to estimate the expected cost J in Eq.(1). For a fully observed dataset, estimation is trivially done by Monte Carlo (MC) integration where we sample from the causal graph in Figure 2 (i.e. run both the agent and the classifier on the dataset) and average the resulting costs. Under observational missingness, however, the available dataset contains missing entries and one cannot sample from $p(X, Y)$, as one only has access to the observed data distribution $p(\bar{R}, \bar{X}, Y)$. We assume observational missingness scenarios without missingness in the label Y and no direct influence of Y on missingness (i.e. no arrow $Y \rightarrow \bar{R}$), but these restrictions can also be relaxed. Note that, in contrast to the AFA missingness mechanism, the observational missingness mechanism is not necessarily restricted to $\mathcal{F}_{MAR}$.

A possible strategy to still run the agent under observational missingness is to allow the agent to acquire only the available features and block the acquisition action if the feature is not available (i.e. if $\bar{R}_i = 0$, then $\hat{R}_i = 0$). This results in a modified causal graph, the *AFA graph* (Figure 3A), which comprises the observational missingness indicator $\bar{R}$ as an environment variable. In Appendix A.3 we provide details about sampling from the AFA graph. The arrow $\bar{R} \rightarrow \hat{R}$ restricts the AFA RL policy to observe only features that are available. The arrow $X \rightarrow \bar{R}$ corresponds to the observational missingness mechanism which is itself represented by an m-graph (as described in Figure 1A).

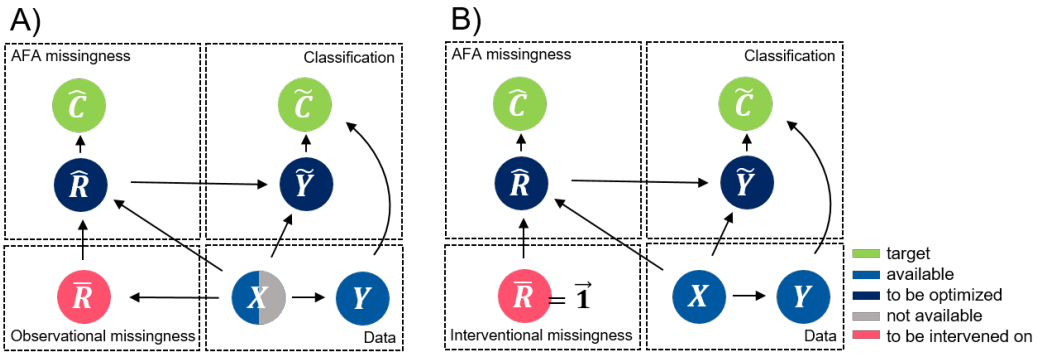

Figure 3: AFA graph: Causal graph for AFA setting with observational missingness. A) Observational environment (for training/testing). B) Interventional environment (at run-time). The difference of the two graphs is characterized by the intervention $do(\bar{R} = \vec{1})$.

Assumption 1, which assumes full control of the AFA agent over the missingness at run-time, implies that the AFA agent is allowed to implement causal interventions on $\bar{R}$. This is contrary to the situation shown in Figure 3A), where $\bar{R}$ depends on X. This causal intervention is mathematically expressed as $do(\bar{R} = \vec{1})$ (Figure 3B). As a result, we define the problem of active feature acquisition performance evaluation (AFAPE) under missing data as the task of evaluating

$$J_{do(\bar{R}=\vec{1})} = \mathbb{E}\left[C \middle| do(\bar{R} = \vec{1})\right]. \tag{2}$$

In the causal graph in Figure 3A), X is a confounder (there exists a path $\bar{R} \leftarrow X \rightarrow C$) which requires adjustment. An exception is MCAR observational missingness which requires no adjustments as there is no arrow $\bar{R} \leftarrow X$.

3.4 SOLUTIONS TO AFAPE

We now examine different views and estimation strategies for the AFAPE problem that make use of different variables in the AFA graph.

3.4.1 BLACK-BOX MISSING DATA VIEW

We call the first view the black-box missing data view. It corresponds to the traditional view on missingness adjustment from the missing data literature, which hasn't been applied in AFA settings previously. It reduces the AFA graph to the three variables $X, \bar{R}$ and C, as shown in Figure 4A). We denote it as 'black-box' since it makes no use of the special structure of the AFA problem (knowledge about $\hat{R}$) and treats the AFA agent and classifier as a black-box function that maps $\bar{R}$ and X to C.

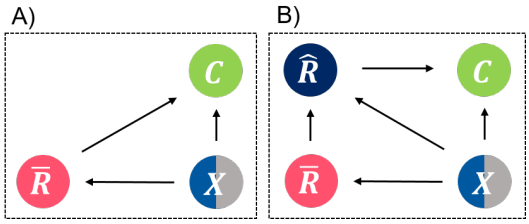

Figure 4: Reduced AFA graph for different estimation strategies. A) Black-box missing data view. B) Off-environment bandit view.

The black-box missing data view leads to the following IS-based estimator for $J_{do(\bar{R}=\vec{1})}$, called inverse probability weighting (IPW) (Seaman & White, 2013):

$$J_{IPW} = \mathbb{E}_{C,X,\bar{R}}\left[C\frac{\mathbb{I}(\bar{R}=\vec{1})}{p(\bar{R}=\vec{1}|X)}\right] = \mathbb{E}_{C,X_{adj},\bar{R}}\left[C\frac{\mathbb{I}(\bar{R}=\vec{1})}{p(\bar{R}=\vec{1}|X_{adj})}\right] \tag{3}$$

where X_{adj} is a subset of features that blocks all backdoor paths between X and C.

Intuition behind IPW:
The importance ratio $\rho_{IPW} = \frac{\mathbb{I}(\bar{R}=\vec{1})}{p(\bar{R}=\vec{1}|X)}$ (where $p(\bar{R}=\vec{1}|X)$ denotes the propensity score) is derived from the importance ratio $\frac{p(X|do(\bar{R}=\vec{1}))}{p(X|\bar{R}=\vec{1})}$. The idea of this IS approach is to sample from the known distribution $p(X|\bar{R}=\vec{1})$ instead of the unknown target feature distribution $p(X|do(\bar{R}=\vec{1}))$, and to adjust for the difference using the so-called importance ratio.

Advantage of IPW:
A big advantage of IPW is its completeness property, i.e. any function of the target distribution is identified if and only if the propensity score is identified (Nabi et al., 2020). Therefore, if identification for a particular missingness scenario fails for this approach, no other approach can perform unbiased estimation.

Disadvantage of IPW:
Under IPW, AFA agents can only be tested on the complete subset of the data (the IS distribution is $p(X|\bar{R}=\vec{1})$). This is a severe disadvantage, as even an agent that never acquires a certain feature could nevertheless be evaluated only on data points where that feature has observed entries. The resulting data inefficiency becomes more significant when the number of complete cases is low, a common scenario in many real-world applications where the number of complete cases can easily be $< 1\%$.

To close the remarks, we point out another common estimation strategy in the missing data literature: multiple imputation (MI) (Sterne et al., 2009). We review this approach for the application on AFA in Appendix A.4. This approach can be adopted to overcome the data inefficiency of IPW. However, it suffers from the disadvantage of a potentially very complicated identification step and the requirement for fitting many conditional densities, which increases the risk of model misspecification. In Appendix A.4, we also discuss common mistakes such as imputing using only features (instead of features and label) and conditional mean imputation, which both can lead to biased results.

3.4.2 OFF-ENVIRONMENT BANDIT VIEW

We now propose a second view for AFAPE that leverages the knowledge about the AFA feature acquisitions $\hat{R}$. It is based on the reduced AFA graph shown in Figure 4B). Although AFA was defined as an RL problem, the order of acquisitions doesn't matter for the evaluation step. We

therefore treat the whole set of acquisitions $\hat{R}$ as one action and the agent as a multiarmed bandit (Auer et al., 2002) instead of a RL agent. Thus, $\bar{R}, X$ and Y can be seen as the environment, while $\hat{R}$ denotes the bandit's action and C the return. The AFAPE problem can thus be seen as off-environment policy evaluation of a bandit under an environment intervention.

Based on the off-environment bandit view, we propose AFAIS (active feature acquisition importance sampling), a novel unbiased and consistent estimator for $J_{do(\bar{R}=\vec{1})}$:

$$J_{AFAIS} = \mathbb{E}_{C,\hat{R},X_{adj},\bar{R}}\left[C\frac{p(\bar{R}=\vec{1}|X_{adj},\hat{R})}{p(\bar{R}=\vec{1}|X_{adj})}\right] \tag{4}$$

The derivation is shown in Appendix A.5.

Intuition behind AFAIS:
The term $\rho_{AFAIS} = \frac{p(\bar{R}=\vec{1}|X_{adj},\hat{R})}{p(\bar{R}=\vec{1}|X_{adj})}$ is derived from the importance ratio $\frac{p(\hat{R}|X_{adj},\bar{R}=\vec{1})}{p(\hat{R}|X_{adj})}$. The idea behind this approach is that we switch the adjustment from different feature distributions (as in IPW) to an adjustment for different AFA feature acquisition distributions. The problem remains that we cannot sample from the policy $p(\hat{R}|X_{adj},\bar{R}=\vec{1})$ for every data point because we do not have certain features (and have to block those actions). Instead, we sample from $p(\hat{R}|X_{adj}) = \sum_{\bar{R},X_{\text{-}adj}} p(\hat{R}|X_{adj},X_{\text{-}adj},\bar{R})p(\bar{R},X_{\text{-}adj}|X_{adj})$ which is possible to do over all data points (not just fully observed data points). The variable set $X_{\text{-}adj}$ corresponds to the subset of features not in X_{adj}.

Advantage of AFAIS:
The obvious and main advantage of AFAIS is that it uses all data points instead of only complete cases (as in IPW) which increases data efficiency. However, this advantage diminishes in case of "data hungry" agents, i.e. agents that often acquire relatively large feature sets. For instance, if the AFA agent always acquires all the features, i.e. $p(\hat{R}=\vec{1}|X_{adj},\bar{R}=\vec{1}) = 1\ \forall X_{adj}$, then the numerator, which we denote as the AFA propensity score, $p(\bar{R}=\vec{1}|X_{adj},\hat{R}) = 0$ if $\hat{R} \neq \vec{1}$ and AFAIS and IPW will have equal weights.

Disadvantage of AFAIS:
One disadvantage of AFAIS is that it requires the fitting of a second function, the AFA propensity score $p(\bar{R}=\vec{1}|X_{adj},\hat{R})$, which increases the chance of model misspecification. Additionally, identification is more demanding for AFAIS. In Appendix A.6, we show the identifiability of J_{AFAIS} under MCAR and simple MAR missingness scenarios, and show how it cannot be evaluated in an MNAR setting.

3.4.3 THE AFAIS-IPW SPECTRUM

AFAIS and IPW can be viewed as two ends of a spectrum. In IPW, we only use the complete cases, while in AFAIS we use all data points with all missingness patterns. A tuning parameter that allows us to gradually move on this spectrum toward each approach, helps us overcome the disadvantages of both (low data efficiency for IPW and problem of evaluation under MNAR missingness for AFAIS). We propose the modified AFAIS estimator as:

$$J_{AFAIS}(s) = \mathbb{E}_{C,\hat{R},X_{adj},\bar{R}}\left[C\frac{p(\bar{R}=\vec{1}|X_{adj},\hat{R},\bar{R}_s=\vec{1})\mathbb{I}(\bar{R}_s=\vec{1})}{p(\bar{R}=\vec{1}|X_{adj})}\right] \tag{5}$$

for any desired index subset $s \subseteq \{0,\dots,d_x\}$. The derivation is shown in Appendix A.7.

Intuition behind the modified AFAIS:
The modified AFAIS importance ratio $\rho_{AFAIS}(s) = \frac{p(\bar{R}=\vec{1}|X_{adj},\hat{R},\bar{R}_s=\vec{1})\mathbb{I}(\bar{R}_s=\vec{1})}{p(\bar{R}=\vec{1}|X_{adj})}$ is derived from the importance ratio $\frac{p(\hat{R}|X_{adj},\bar{R}=\vec{1})}{p(\hat{R}|X_{adj},\bar{R}_s=\vec{1})}\frac{p(X_{adj}|do(\bar{R}=\vec{1}))}{p(X_{adj}|\bar{R}_s=\vec{1})}$. The idea behind this approach is to sample only from datapoints where an important set of features (the features in s) are always observed (i.e. the feature distribution $p(X_{adj}|\bar{R}_s=\vec{1})$) while using the off-environment bandit approach for the remaining features. To showcase the advantages of using the modified AFAIS, we demonstrate in Appendix A.8 that the modified AFAIS is identified for the MNAR example from Appendix A.6 and that it can be evaluated when an appropriate conditioning set s is chosen.

3.4.4 IDENTIFICATION

We do not give general identifiability statements for the modified AFAIS estimator, but suggest the following guide for choosing the set s that yields the most data efficient scenario under which the modified AFAIS can be used: First, start with $s = \emptyset$ (full off-environment bandit view) and investigate identifiability and the possibility of evaluation. If identification and evaluation is possible, then the entire dataset can be used for estimation. Otherwise, try larger sets for s until identification and evaluation is possible or $s = \{0, \dots, d_x\}$ has been reached. When the set $s = \{0, \dots, d_x\}$ is reached (full black-box missing data view), then $J_{AFAIS}(s) \equiv J_{IPW}$ and the IPW known identification theories can be utilized.

3.4.5 OTHER VIEWS ON SOLVING AFAPE

We also would like to mention other views on solving AFAPE. These include off-policy policy evaluation (OPE) which has been used in the AFA literature before, but is very data inefficient and can lead to biased results in complex missing data scenarios. Another approach used in the AFA literature is to just block the acquisitions without adjustment which thus also leads to biased results. We review these two approaches in Appendix A.9.

4 EXPERIMENTS

4.1 EXPERIMENT DESIGN

We evaluate different estimators on synthetic and real world datasets under synthetic MCAR, MAR and MNAR observational missingness. We evaluate random AFA policies, CATE (Das et al., 2021) and a vanilla Deep Q-Network (DQN) RL agent (Mnih et al., 2015) as AFA agents and use a random forest classifier. We compare the following estimators:

- J_{Imp} uses Missforest (Stekhoven & Bühlmann, 2012) to impute missing features based on observed features.
- $J_{\bar{R}}$ blocks the acquisitions of not available features, but offers no correction.
- $J_{\bar{R}=\vec{1}}$ corresponds to complete case analysis and is only unbiased under MCAR.
- J_{IPW} is the stabilized IPW estimator. Stabilization is achieved by normalizing the weights which reduces variance at the cost of small bias (Rubinstein & Kroese, 2016).
- J_{AFAIS} is the stabilized (modified) AFAIS estimator.
- $J_{do(\bar{R}=\vec{1})}$ is the ground truth, where the agent is run on the fully observed dataset.

We fit the propensity score and AFA propensity score using fully-connected neural networks. Complete experiment details are given in Appendix A.10.

4.2 RESULTS

Figure 5 depicts convergence rates of different estimators for the MNAR synthetic data experiment for two random policies as a function of evaluation sample size. The results show that the three knowingly-biased estimators J_{Imp}, $J_{\bar{R}}$ and $J_{\bar{R}=\vec{1}}$ do not converge to correct solutions as the sample size increases. The stabilized AFAIS (and IPW) estimators do converge to the correct solution, but at different rates depending on the policy. In the case of a policy that only acquires 10% of the costly features, the stabilized AFAIS estimator J_{AFAIS} is converging almost as fast as the ground truth $J_{do(\bar{R}=\vec{1})}$ (left plot). In the case where the agent acquires 90% of the data, the AFAIS and IPW estimators are converging equally fast and there is no benefit of using AFAIS over the conventional IPW estimator, as expected (right plot).
Figure 6 shows the resulting estimates for $J_{do(\bar{R}=\vec{1})}$ for the MNAR synthetic data experiment for different agents. Only J_{IPW} and J_{AFAIS} show unbiased results. The experiment demonstrates the large errors biased estimators, such as J_{Imp}, can produce and the potentially wrong conclusions these estimators can lead to: If the policies are compared according to J_{Imp}, an agent that acquires each costly feature with 10% probability is to be preferred over an agent with random 90% acquisition.

At run-time, the 10% agent would perform, however, significantly worse (higher $J_{do(\bar{R}=\vec{1})}$). Further real world data experiment results are shown in Appendix A.11.

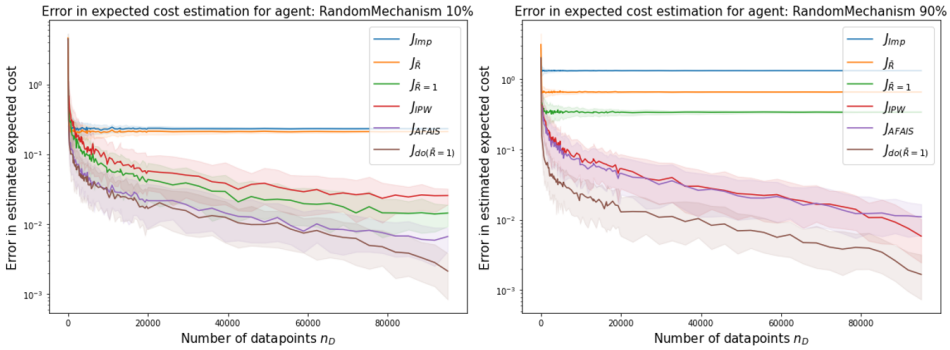

Figure 5: Error of expected cost estimation as a function of the number of data points in a synthetic data experiments with MNAR missingness. Plot shows two random agents that acquire each costly feature with a probability of 10% (left) and 90% (right). Error is defined as the difference between an estimator and the ground truth estimator $J_{do(\bar{R}=\vec{1})}$ evaluated on the whole dataset $(n_D = 100000)$.

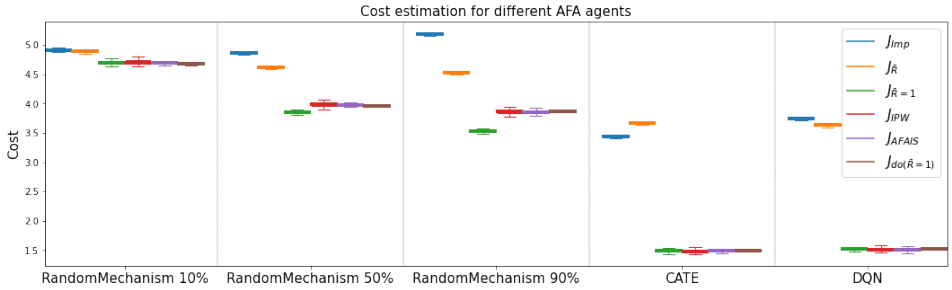

Figure 6: Evaluated performance of different agents for the MNAR synthetic data experiment. Box plots shows variability of the estimate by bootstrapping 100 times from the total 100000 MC samples.

5 Discussion and Future Work

The central argument of this paper is based on Assumption 1, the full control of the AFA agent over the missingness at run-time. This assumption might, however, not always hold in practice (e.g. due to the breaking of a sensor or a patient who is not cooperative). In this case, a different distribution shift for $\bar{R}$ should be assumed for evaluation.

We further consider the following suggestions as major extensions of this work: i) studying the evaluation of AFA agents under missingness in the temporal setting with time-series features; ii) exploring other prediction tasks (other than classification); iii) including the counterfactual expected cost $J_{do(\bar{R}=\vec{1})}$ in the training of AFA agents and classifiers.

6 Conclusion

We study the problem of active feature acquisition performance evaluation (AFAPE) under missing data which corresponds to an intervention on the RL environment used to train AFA agents. We propose a new causal graph, the AFA graph, and show that current evaluation methods for the AFAPE problem can be strongly biased leading to high risks for model deployment. AFAIS, a novel IS-based estimator, gives unbiased solutions under MCAR, MAR and MNAR missingness at a higher data efficiency than IPW.

REFERENCES

Chaojie An, Qifeng Zhou, and Shen Yang. A reinforcement learning guided adaptive cost-sensitive feature acquisition method. *Applied Soft Computing*, pp. 108437, January 2022. ISSN 15684946. doi: 10.1016/j.asoc.2022.108437. URL https://linkinghub.elsevier.com/retrieve/pii/S1568494622000163.

Qi An, Hui Li, Xuejun Liao, and Lawrence Carin. Active feature acquisition with POMDP models. *Submitted to Pattern Recognition Letters*, 2006.

Bálint Antal and András Hajdu. An ensemble-based system for automatic screening of diabetic retinopathy. *Knowledge-based systems*, 60:20–27, 2014.

Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The Nonstochastic Multi-armed Bandit Problem. *SIAM Journal on Computing*, 32(1):48–77, January 2002. ISSN 0097-5397. doi: 10.1137/S0097539701398375. URL https://epubs.siam.org/doi/abs/10.1137/S0097539701398375. Publisher: Society for Industrial and Applied Mathematics.

Christian Beyer, Maik Büttner, Vishnu Unnikrishnan, Miro Schleicher, Eirini Ntoutsi, and Myra Spiliopoulou. Active feature acquisition on data streams under feature drift. *Annals of Telecommunications*, 75(9-10):597–611, October 2020. ISSN 0003-4347, 1958-9395. doi: 10.1007/s12243-020-00775-2. URL https://link.springer.com/10.1007/s12243-020-00775-2.

Rohit Bhattacharya, Razieh Nabi, Ilya Shpitser, and James M. Robins. Identification In Missing Data Models Represented By Directed Acyclic Graphs. In *Proceedings of The 35th Uncertainty in Artificial Intelligence Conference*, pp. 1149–1158. PMLR, August 2020. URL https://proceedings.mlr.press/v115/bhattacharya20b.html.

Chun-Hao Chang, Mingjie Mai, and Anna Goldenberg. Dynamic Measurement Scheduling for Event Forecasting using Deep RL. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 951–960. PMLR, May 2019. URL https://proceedings.mlr.press/v97/chang19a.html.

Srijita Das, Rishabh Iyer, and Sriraam Natarajan. A Clustering based Selection Framework for Cost Aware and Test-time Feature Elicitation. In *8th ACM IKDD CODS and 26th COMAD*, pp. 20–28. ACM, January 2021. ISBN 978-1-4503-8817-7. doi: 10.1145/3430984.3431008. URL https://dl.acm.org/doi/10.1145/3430984.3431008.

Miroslav Dudik, John Langford, and Lihong Li. Doubly Robust Policy Evaluation and Learning, May 2011. URL http://arxiv.org/abs/1103.4601. arXiv:1103.4601 [cs, stat].

Gabriel Erion, Joseph D. Janizek, Carly Hudelson, Richard B. Utarnachitt, Andrew M. McCoy, Michael R. Sayre, Nathan J. White, and Su-In Lee. CoAI: Cost-Aware Artificial Intelligence for Health Care. Technical report, medRxiv, January 2021. URL https://www.medrxiv.org/content/10.1101/2021.01.19.21249356v1.

FICO. FICO Explainable Machine Learning Challenge. https://community.fico.com/s/explainable-machine-learning-challenge, May 2018. URL https://community.fico.com/s/explainable-machine-learning-challenge. Last accessed May 2022.

Wenbo Gong, Sebastian Tschiatschek, Sebastian Nowozin, Richard E Turner, José Miguel Hernández-Lobato, and Cheng Zhang. Icebreaker: Element-wise Efficient Information Acquisition with a Bayesian Deep Latent Gaussian Model. In *Advances in Neural Information Processing Systems*, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/c055dcc749c2632fd4dd806301f05ba6-Abstract.html.

Sheng-Jun Huang, Miao Xu, Ming-Kun Xie, Masashi Sugiyama, Gang Niu, and Songcan Chen. Active Feature Acquisition with Supervised Matrix Completion. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1571–1579, July 2018. ISBN 978-1-4503-5552-0. doi: 10.1145/3219819.3220084. URL https://dl.acm.org/doi/10.1145/3219819.3220084.

Jaromír Janisch, Tomáš Pevný, and Viliam Lisý. Classification with costly features as a sequential decision-making problem. *Machine Learning*, 109(8):1587–1615, August 2020. ISSN 0885-6125, 1573-0565. doi: 10.1007/s10994-020-05874-8. URL http://link.springer.com/10.1007/s10994-020-05874-8.

Daniel Jarrett and Mihaela van der Schaar. Inverse Active Sensing: Modeling and Understanding Timely Decision-Making. *arXiv:2006.14141 [cs, stat]*, June 2020. URL http://arxiv.org/abs/2006.14141.

Nathan Kallus and Masatoshi Uehara. Double Reinforcement Learning for Efficient Off-Policy Evaluation in Markov Decision Processes. *Journal of Machine Learning Research*, 21(167):1–63, 2020. ISSN 1533-7928. URL http://jmlr.org/papers/v21/19-827.html.

Thomas T. Kok, Rachel M. Brouwer, Rene M. Mandl, Hugo G. Schnack, and Georg Krempl. Active Selection of Classification Features. *arXiv:2102.13636 [cs]*, February 2021. URL http://arxiv.org/abs/2102.13636. arXiv: 2102.13636.

Marine Le Morvan, Julie Josse, Erwan Scornet, and Gael Varoquaux. What's a good imputation to predict with missing values? In *Advances in Neural Information Processing Systems*, volume 34, pp. 11530–11540, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/5fe8fdc79ce292c39c5f209d734b7206-Abstract.html.

Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu. Offline Reinforcement Learning: Tutorial, Review, and Perspectives on Open Problems. *arXiv:2005.01643 [cs, stat]*, November 2020. URL http://arxiv.org/abs/2005.01643. arXiv: 2005.01643.

Yang Li and Junier Oliva. Active Feature Acquisition with Generative Surrogate Models. In *Proceedings of the 38th International Conference on Machine Learning*, pp. 6450–6459. PMLR, July 2021a. URL https://proceedings.mlr.press/v139/li21p.html. ISSN: 2640-3498.

Yang Li and Junier B. Oliva. Dynamic Feature Acquisition with Arbitrary Conditional Flows. *arXiv:2006.07701 [cs, stat]*, March 2021b. URL http://arxiv.org/abs/2006.07701.

Yang Li, Siyuan Shan, Qin Liu, and Junier B. Oliva. Towards Robust Active Feature Acquisition. *arXiv:2107.04163 [cs]*, July 2021. URL http://arxiv.org/abs/2107.04163.

Charles X. Ling, Qiang Yang, Jianning Wang, and Shichao Zhang. Decision trees with minimal costs. In *Twenty-first international conference on Machine learning - ICML '04*, pp. 69, 2004. doi: 10.1145/1015330.1015369. URL http://portal.acm.org/citation.cfm?doid=1015330.1015369.

Roderick JA Little and Donald B. Rubin. *Statistical analysis with missing data*, volume 793. John Wiley & Sons, 2019.

Chao Ma, Sebastian Tschiatschek, Konstantina Palla, Jose Miguel Hernandez-Lobato, Sebastian Nowozin, and Cheng Zhang. EDDI: Efficient Dynamic Discovery of High-Value Information with Partial VAE. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 4234–4243. PMLR, May 2019. URL https://proceedings.mlr.press/v97/ma19c.html.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, February 2015. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature14236. URL http://www.nature.com/articles/nature14236.

Karthika Mohan, Judea Pearl, and Jin Tian. Graphical Models for Inference with Missing Data. In *Advances in Neural Information Processing Systems*, volume 26, 2013. URL https://proceedings.neurips.cc/paper/2013/hash/0ff8033cf9437c213ee13937b1c4c455-Abstract.html.

Rémi Munos and Csaba Szepesvári. Finite-Time Bounds for Fitted Value Iteration. *Journal of Machine Learning Research*, 9(5), 2008.

Razieh Nabi, Rohit Bhattacharya, and Ilya Shpitser. Full Law Identification In Graphical Models Of Missing Data: Completeness Results. *arXiv:2004.04872 [cs, stat]*, August 2020. URL http://arxiv.org/abs/2004.04872.

Sriraam Natarajan, Srijita Das, Nandini Ramanan, Gautam Kunapuli, and Predrag Radivojac. On Whom Should I Perform this Lab Test Next? An Active Feature Elicitation Approach. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pp. 3498–3505, Stockholm, Sweden, July 2018. International Joint Conferences on Artificial Intelligence Organization. ISBN 978-0-9992411-2-7. doi: 10.24963/ijcai.2018/486. URL https://www.ijcai.org/proceedings/2018/486.

D. J. Newman, S. Hettich, C. L. Blake, and C. J. Merz. UCI Repository of machine learning databases, 1998. URL http://www.ics.uci.edu/~mlearn/MLRepository.html.

Judea Pearl. *Causality*. Cambridge university press, 2009. ISBN 978-0-521-89560-6.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Doina Precup, Richard S. Sutton, and Satinder Singh. Eligibility Traces for Off-Policy Policy Evaluation, 2000.

Reuven Y Rubinstein and Dirk P Kroese. *Simulation and the Monte Carlo method*. John Wiley & Sons, 2016.

Shaun R. Seaman and Ian R. White. Review of inverse probability weighting for dealing with missing data. *Statistical methods in medical research*, 22(3):278–295, 2013. Publisher: Sage Publications Sage UK: London, England.

Burr Settles. Active learning literature survey. 2009.

Victor S. Sheng and Charles X. Ling. Feature value acquisition in testing: a sequential batch test algorithm. In *Proceedings of the 23rd international conference on Machine learning*, pp. 809–816, 2006.

Hajin Shim, Sung Ju Hwang, and Eunho Yang. Joint Active Feature Acquisition and Classification with Variable-Size Set Encoding. *Advances in Neural Information Processing Systems*, 31, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/e5841df2166dd424a57127423d276bbe-Abstract.html.

Ilya Shpitser, Karthika Mohan, and Judea Pearl. Missing data as a causal and probabilistic problem. In *Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence*, UAI'15, pp. 802–811, Arlington, Virginia, USA, July 2015. AUAI Press. ISBN 978-0-9966431-0-8.

Daniel J. Stekhoven and Peter Bühlmann. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1):112–118, 2012.

Jonathan A. C. Sterne, Ian R. White, John B. Carlin, Michael Spratt, Patrick Royston, Michael G. Kenward, Angela M. Wood, and James R. Carpenter. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, 338:b2393, June 2009. ISSN 0959-8138, 1468-5833. doi: 10.1136/bmj.b2393. URL https://www.bmj.com/content/338/bmj.b2393. Publisher: British Medical Journal Publishing Group Section: Research Methods & Reporting.

Fengyi Tang, Lifan Zeng, Fei Wang, and Jiayu Zhou. Adversarial Precision Sensing with Healthcare Applications. In *2020 IEEE International Conference on Data Mining (ICDM)*, pp. 521–530, November 2020. doi: 10.1109/ICDM50108.2020.00061.

Stef van Buuren. Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, 16(3):219–242, June 2007. ISSN 0962-2802. doi: 10.1177/0962280206074463. URL https://doi.org/10.1177/0962280206074463.

Thomas S. Verma and Judea Pearl. Equivalence and Synthesis of Causal Models. Technical Report R-150, Department of Computer Science, University of California, Los Angeles, 1990.

Xiaoyong Chai, Lin Deng, Qiang Yang, and C. X. Ling. Test-cost sensitive naive Bayes classification. In *Fourth IEEE International Conference on Data Mining (ICDM'04)*, pp. 51–58, November 2004. doi: 10.1109/ICDM.2004.10092.

Haiyan Yin, Yingzhen Li, Sinno Jialin Pan, Cheng Zhang, and Sebastian Tschiatschek. Reinforcement Learning with Efficient Active Feature Acquisition. *arXiv:2011.00825 [cs]*, November 2020. URL http://arxiv.org/abs/2011.00825.

Jinsung Yoon, William R. Zame, and Mihaela Van Der Schaar. Deep sensing: Active sensing using multi-directional recurrent neural networks. In *International Conference on Learning Representations*, 2018.

Jinsung Yoon, James Jordon, and Mihaela Schaar. ASAC: Active Sensing using Actor-Critic models. In *Machine Learning for Healthcare Conference*, pp. 451–473. PMLR, October 2019. URL http://proceedings.mlr.press/v106/yoon19a.html. ISSN: 2640-3498.

Pin Zhang. A novel feature selection method based on global sensitivity analysis with application in machine learning-based prediction model. *Applied Soft Computing*, 85:105859, 2019. Publisher: Elsevier.

A Appendix

A.1 Literature review for active feature acquisition

Active feature acquisition (AFA) (An et al., 2006; Li & Oliva, 2021a; Li et al., 2021; Chang et al., 2019; Shim et al., 2018; Yin et al., 2020) has been studied under various other names including, but not limited to, "active sensing" (Yoon et al., 2019; 2018; Tang et al., 2020; Jarrett & van der Schaar, 2020), "active feature elicitation" (Natarajan et al., 2018; Das et al., 2021),"dynamic feature acquisition" (Li & Oliva, 2021b), "dynamic active feature selection" (Zhang, 2019), "element-wise efficient information acquisition" (Gong et al., 2019), "classification with costly features" (Janisch et al., 2020) and "test-cost sensitive classification" (Xiaoyong Chai et al., 2004). AFA is different from active learning (Settles, 2009). In active learning, one assumes a classification task with a training data set that contains many unlabeled data points. The active learning task is then to decide which label acquisitions will improve the training performance the most. Similar research also exists for the acquisition of features for optimal improvement of training. This task has been referred to as "active selection of classification features" (Kok et al., 2021), and unfortunately also "active feature acquisition" (Huang et al., 2018; Beyer et al., 2020), but its objective differs fundamentally from ours. Huang et al. (2018) attempt to find out which missing values within the retrospective data set would improve training the most when retroactively acquired. In this paper, we are, however, interested which features, for a new data point, would improve the individual prediction for that data point the most.

Many AFA approaches are based on a greedy feature acquisition strategy wrapping a subsequent classification task. An idea is to employ decision tree classifiers and to acquire features sequentially by traversing the branch of the decision tree (Ling et al., 2004; Sheng & Ling, 2006), while the splitting criteria minimizes the combined cost of feature acquisition and misclassification (Ling et al., 2004). Another approach, the test-cost sensitive Naive Bayes (csNB) classifier (Xiaoyong Chai et al., 2004), exploits the Naive Bayes assumption of independence among the predictive power of features. This allows for an efficient exploration whether acquiring a certain feature can reduce costs. Das et al. (2021) propose clustering-based cost-aware feature elicitation (CATE). In CATE, data points are clustered based on a set of zero-cost features and the optimal, fixed set of features is computed for each cluster. A new partially-observed data point is then attributed to a cluster, and the corresponding optimal feature set is acquired for it.

The AFA problem is inherently a sequential decision process which one can address using reinforcement learning (RL). Model-based RL approaches learn the state-transition function which, for AFA, corresponds to learning an imputation model (Yoon et al., 2018; Yin et al., 2020; Li & Oliva, 2021a;b; Ma et al., 2019). The imputation model is used at run-time to simulate possible outcomes of a feature acquisition and derive desired acquisition strategies. Alternatively, model-free RL approaches do not require learning a state-transition function. One variant, Q-learning, relies on modeling the expected cost of a particular acquisition decision (Chang et al., 2019; Janisch et al., 2020; Shim et al., 2018). As an example, Shim et al. (2018) use double Q-learning for the AFA agent with a deep neural network that shares network layers with the subsequent classification neural network.

A.2 Missing data models

Missing data problems relate the full data distribution $p(X)$ and the observed data distribution where variables in X may sometimes be unobserved. Each $X_i \in X$ is observed if the corresponding missingness indicator $\bar{R}_i$ assumes value 1. Thus, the observed version of X_i which we denote by $\bar{X}_i$ is defined to be equal to X_i if $\bar{R}_i = 1$, and is missing (equal to a special value "?") otherwise. We further consider an acquisition cost $\bar{C}_i$ for the feature X_i. The cost $\bar{C}$ is thus deterministically defined by the missingness indicator $\bar{R}$. The goal in missing data problems is to estimate some function of $p(X)$, such as a parameter or a predictive model, using the i.i.d. samples from the observed distribution $p(\bar{X}, \bar{R})$, where $\bar{X}$ is the set of all observed proxies, and $\bar{R}$ is the set of all missingness indicators in the problem. Parameters of interest must be identified, i.e. unique functions of $p(\bar{X}, \bar{R})$, in order to make the estimation problem well-posed.

The problem of identification is to determine whether or not a function of the full data can be recovered from the observed data distribution. Identification is not possible in general, and relies on assumptions encoded in a missing data model. The simplest identified missing data model is the missing completely at random (MCAR) model, which assumes that $p(\bar{R}|X) = p(\bar{R})$. The missing at random model (MAR) assumes that missingness only depends on observed features: $p(\bar{R} = \bar{r}|X) = p(\bar{R} = \bar{r}|\{X_i : \bar{r}_i = 1\})$. Finally, any model that is neither MCAR nor MAR is called missing not at random (MNAR) where missingness can now depend also on feature values that are unobserved. While parameters of interest are not identified in general under MNAR, a large class of identified MNAR models has been derived in the missing data literature (Bhattacharya et al., 2020; Nabi et al., 2020).

In this paper, we will restrict attention to a special type of missing data models where identifiability restrictions are represented by a directed acyclic graph (DAG) factorization of the distribution. The restrictions are then formalized as $p(X, \bar{R}) = \prod_{V \in X \cup \bar{R}} p(V|\text{pa}_{\mathcal{G}}(V))$ for some graph $\mathcal{G}$ where $\text{pa}_{\mathcal{G}}(V)$ selects the parents of the node V in $\mathcal{G}$. The graph $\mathcal{G}$ is termed the *missing data graph* (m-graph) (Mohan et al., 2013; Shpitser et al., 2015) (Figure 1A). In models of missing data, the identification problem for parameters in $p(X)$ from $p(\bar{X}, \bar{R})$ may be formulated as an identification problem of causal inference, where $p(X)$ is obtained as a result of applying the intervention operator do(.) (Pearl, 2009) to every missingness indicator $\bar{R}$, hence $p(X) = p(\bar{X}|\text{do}(\bar{R} = \vec{1})) = \frac{p(\bar{X}, \bar{R} = \vec{1})}{p(\bar{R} = \vec{1}|X)}$. In other words, the problem of identification of $p(X)$ is translated to the problem of identification of $p(\bar{R} = \vec{1}|X)$. If the missing data model is graphical, techniques from causal effect identification theory may be applied to obtain identification of $p(X)$ (Bhattacharya et al., 2020; Nabi et al., 2020).

A.3 Sampling from the AFA graph

In this section, we provide the details of factorizing the joint distribution over the variables of the AFA process with respect to the AFA graph (Figure 3) into distributions from which we can sample. The joint distribution factorizes as

$$p(\tilde{C}, \tilde{Y}, \hat{C}, \hat{R}, \bar{R}, X, Y) = p(\tilde{C}|\tilde{Y}, Y)p(\hat{C}|\hat{R})p(\tilde{Y}|X, \hat{R})p(\hat{R}|X, \bar{R})p(\bar{R}, X, Y). \tag{A.1}$$

The sampling procedure thus is defined as follows. First, we sample from the retrospective dataset $p(\bar{R}, X, Y)$ (i.e. the "environment"). Next, we run the AFA agent. This gives us samples from the AFA missingness $p(\hat{R}|X, \bar{R})$ and the (deterministic) acquisition cost $p(\hat{C}|\hat{R})$ conditional distributions. As AFA missingness is not modelled from a set viewpoint (i.e. we cannot directly sample $p(\hat{R}|X, \bar{R})$), we sample equivalently from the corresponding AFA RL policy $\pi(\hat{U}|\hat{X}^{(k)}, \hat{R}^{(k)}, \bar{R})$.

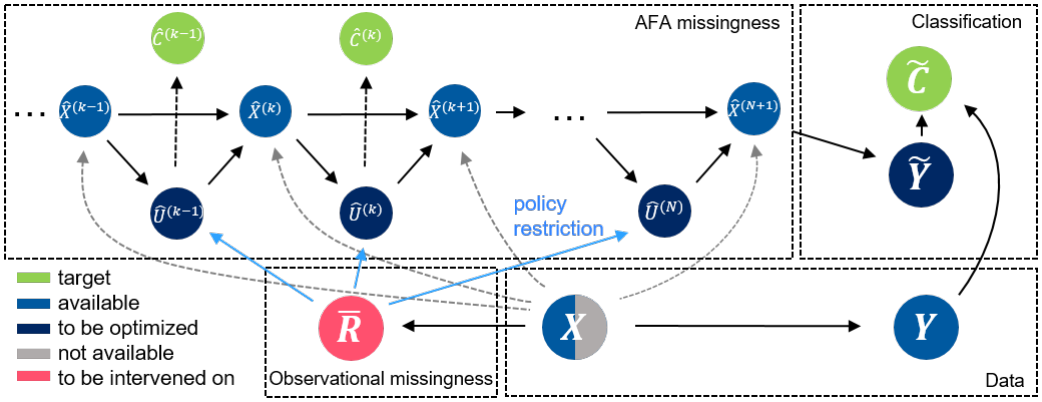

Figure A.1: AFA graph from an AFA missingness policy viewpoint. The node $\hat{X}^{(k)}$ represents the intermediate states at step k of the decision process. The state also contains $\hat{R}^{(k)}$ which was left out for clarity). The AFA policy $\pi(\bar{U}^{(k)}|\hat{X}^{(k)}, \hat{R}^{(k)}, \bar{R})$ is restricted by $\bar{R}$ to only allow the acquisition of available features, i.e. $\pi(\bar{U}^{(k)} =$ "observe(X_i)"$|\hat{X}^{(k)}, \hat{R}^{(k)}, \bar{R} = \bar{r}) = 0$ if $\bar{r}_i = 0$.

To visualize this process, we present in Figure A.1 the AFA graph from an AFA missingness sequence perspective. Sampling $\hat{R}$ now corresponds to sampling $\hat{U}$. We block feature acquisition actions if the desired feature is not available (i.e. $p(\hat{R} = \hat{r}|X, \bar{R} = \bar{r}) = 0$ if $\hat{r}_i = 1$ and $\bar{r}_i = 0$). Finally, we apply the (deterministic) classifier and compute the misclassification cost (i.e. "sample" from $p(\tilde{Y}|X, \hat{R})$ and $p(\tilde{C}|\tilde{Y}, Y)$).

A.4 Multiple Imputation (MI) for the AFAPE problem

An alternative estimation strategy to IPW based on the missing data view is multiple imputation (MI). The core idea of imputation is to fill in (impute) missing entries and then assume a fully observed dataset. The procedure can be formulated as (Little & Rubin, 2019)

$$\begin{aligned} J_{do(\bar{R}=\vec{1})} = \mathbb{E}_{X,Y}[h(X,Y)] &= \mathbb{E}_{\bar{X},Y}\left[\mathbb{E}_{X_{miss}}[h(\bar{X}, X_{miss}, Y)]\right] \\ &= \sum_{\bar{X},Y}\sum_{X_{miss}} h(\bar{X}, X_{miss}, Y)p(X_{miss}|\bar{X}, Y)p(\bar{X}, Y) \end{aligned} \tag{A.2}$$

where $h(X, Y) \equiv C$ is the black-box function mapping the data to the costs and X_{miss} represents the unobserved data. Imputation thus corresponds to Monte Carlo integration by sampling observed values $p(\bar{X}, Y)$ and imputing missing values using an imputation model $p(X_{miss}|\bar{X}, Y)$. To increase accuracy of the estimator, the imputation can be repeated multiple times leading to the MI procedure. In the identification step, we determine whether the imputation model can be estimated from observational data.

The first drawback of imputation is the requirement of modeling joint distributions which is in practice a complex problem, especially in high-dimensional settings and under complex missingness patterns. As an example, the multiple imputation by chained equations (MICE) method (van Buuren, 2007) requires fitting K conditional densities for K partially observed features. By comparison, IPW requires only propensity score, which is often much easier to specify.

Eq.(A.2) implies imputation by conditioning on the features X and the label Y. This introduces the risk of data leakage, as the imputed features may carry predictive information not because of the true data generation mechanism, but because of the imputation, resulting in a potentially overoptimistic estimation of prediction performance. A common alternative, often used in ML, is to impute the data without conditioning on Y. This assumption, however, implies that a missing feature $X_{\mathrm{i}} \in X_{miss}$ is conditionally independent of the label given the observed features ($X_{\mathrm{i}} \perp\!\!\!\perp Y|\bar{X}$). Determining marginal predictive value of a feature for predicting Y, is however, the whole task of AFA, which renders this approach impractical.

A.5 DERIVATION OF ACTIVE FEATURE ACQUISITION IMPORTANCE SAMPLING (AFAIS)

The AFAIS estimator is based on IS, a popular method in OPE and missing data to adjust for distribution shift. IS assumes a problem setting where one has only access to samples from a "behavior" distribution, while the goal is the evaluation of a "target" distribution. This is made possible when knowledge of the ratio of the two distributions (importance ratio) is available. In the setting of this paper, the target distribution is the distribution of C under the intervention $do(\bar{R} = \vec{1})$. The derivation of AFAIS is based on the reduced AFA graph from Figure 4B) which results from a latent projection (Verma & Pearl, 1990) of the full AFA graph from Figure 3A). This reduces the graph factorization from equation Eq. (A.1) to only the features C, $\hat{R}$, $\bar{R}$ and X. We further consider only a subset of the features X_{adj} that blocks all backdoor paths from $\bar{R}$ to $\hat{R}$ and C. Since this yields a causal model of a DAG, we can apply the truncated factorization, or the g-formula, to obtain:

$$p(C, \hat{R}, X_{adj}|do(\bar{R} = \vec{1})) = p(C|X_{adj}, \hat{R})p(\hat{R}|X_{adj}, \bar{R} = \vec{1})p(X_{adj}) \tag{A.3}$$

We propose the following behavior distribution:

$$p(C, \hat{R}, X_{adj}) = p(C|X_{adj}, \hat{R})p(\hat{R}|X_{adj})p(X_{adj}) \tag{A.4}$$

This means we sample from all data points and not only from observed datapoints as for IPW. Sampling from Eq. (A.4) is equivalent to sampling from $p(C, \hat{R}, X_{adj}, X_{\text{-}adj}, \bar{R}) = p(C, \hat{R}, X, \bar{R}) = p(C|X, \hat{R})p(\hat{R}|X, \bar{R})p(X, \bar{R})$ which one can do as shown in Appendix A.3.

Based on the introduced target and behavior distribution, we now derive the AFAIS estimator:

$$\begin{aligned}
J_{do(\bar{R}=\vec{1})} &= \mathbb{E}\left[C\middle|do(\bar{R} = \vec{1})\right] = \sum_{C} Cp(C|do(\bar{R} = \vec{1})) \\
&= \sum_{C, X_{adj}, \hat{R}} Cp(C, X_{adj}, \hat{R}|do(\bar{R} = \vec{1})) \\
&= \sum_{C, X_{adj}, \hat{R}} Cp(C|X_{adj}, \hat{R})p(\hat{R}|X_{adj}, \bar{R} = \vec{1})p(X_{adj}) \\
&= \sum_{C, X_{adj}, \hat{R}} Cp(C|X_{adj}, \hat{R})p(\hat{R}|X_{adj})p(X_{adj}) \cdot \frac{p(\hat{R}|X_{adj}, \bar{R} = \vec{1})}{p(\hat{R}|X_{adj})} \\
&\overset{(*)}{=} \sum_{C, X_{adj}, \hat{R}} Cp(C|X_{adj}, \hat{R})p(\hat{R}|X_{adj})p(X_{adj}) \cdot \frac{p(\bar{R} = \vec{1}|X_{adj}, \hat{R})}{p(\bar{R} = \vec{1}|X_{adj})} \\
&= \mathbb{E}_{C, \hat{R}, X_{adj}, \bar{R}}\left[C\frac{p(\bar{R} = \vec{1}|X_{adj}, \hat{R})}{p(\bar{R} = \vec{1}|X_{adj})}\right] \equiv J_{AFAIS}
\end{aligned} \tag{A.5}$$

where the equivalence $(*)$ holds due to Bayes rule:

$$\frac{p(\hat{R}|X_{adj}, \bar{R} = \vec{1})}{p(\hat{R}|X_{adj})} = p(\hat{R}|X_{adj}, \bar{R} = \vec{1})\frac{p(\bar{R} = \vec{1}|X_{adj}, \hat{R})}{p(\hat{R}|X_{adj}, \bar{R} = \vec{1})p(\bar{R} = \vec{1}|X_{adj})} = \frac{p(\bar{R} = \vec{1}|X_{adj}, \hat{R})}{p(\bar{R} = \vec{1}|X_{adj})}$$

A.6 IDENTIFICATION EXAMPLE FOR AFAIS

We give an instructive example of how to perform identification for the propensity score $p(\bar{R} = \vec{1}|X_{adj})$ and AFA propensity score $p(\bar{R} = \vec{1}|X_{adj}, \hat{R})$, needed for the AFAIS estimator J_{AFAIS}. We show that identification is possible for a simple MAR (and therefore MCAR) scenario, but that the evaluation of the AFAIS estimator is not possible for an MNAR missing data scenario.

MAR

We assume the missing data graph in Figure A.2A) showing both observational and AFA missingness. The figure shows a simple MAR observational missing data scenario, where X_1 and X_2 are

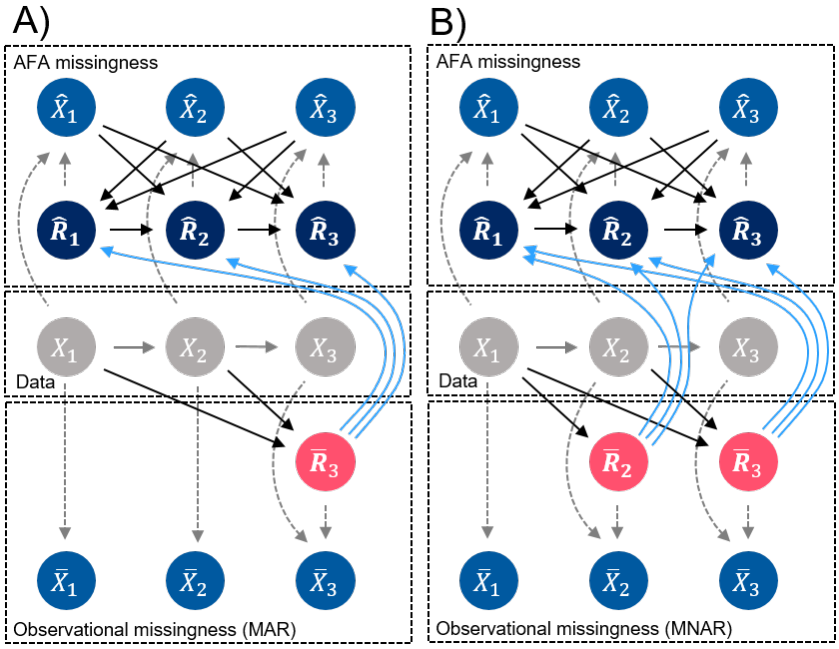

Figure A.2: Missing data graph including the AFA and observational missingness. A) Observational missingness is MAR. B) Observational missingness is MNAR.

not missing in the observational dataset (we leave out $\bar{R}_1$ and $\bar{R}_2$ for clarity), but influence observational missingness in X_3 ($X_1 \rightarrow \bar{R}_3$; $X_2 \rightarrow \bar{R}_3$). The blocking of actions leads to: $\bar{R}_3 \rightarrow \hat{R}_i$ $\forall i$. Furthermore, the AFA agent's decision to acquire a feature can be made based on all previously acquired features which translates to $\hat{X}_i \rightarrow \hat{R}_j$ $\forall i \neq j$. Figure A.2 shows arrows $\hat{R}_1 \rightarrow \hat{R}_2$ and $\hat{R}_2 \rightarrow \hat{R}_3$, but this does not hold in general. Other causes between missingness indicators are in general possible, but this does not affect the identification procedure described in the following. The corresponding adjustment set is $X_{adj} = \{X_1, X_2\}$.

Identification of $p(\bar{R} = \vec{1}|X_{adj})$:
The propensity score is identified by:

$$p(\bar{R} = \vec{1}|X_{adj}) = p(\bar{R}_3 = 1|X_1, X_2)$$

which is a function of the observed data (X_1 and X_2 are always observed).

Identification of $p(\bar{R} = \vec{1}|X_{adj}, \hat{R})$:
The AFA propensity score is also identified:

$$p(\bar{R} = \vec{1}|X_{adj}, \hat{R}) = p(\bar{R}_3 = 1|X_1, X_2, \hat{R})$$

MNAR

We now assume the missing data graph in Figure A.2B). The missingness scenario is equal to the previously described MAR case, except for an additional missingness in $\bar{R}_2$ caused by X_1 ($X_1 \rightarrow \bar{R}_2$). As the missingness indicator $\bar{R}_3$ now depends on a potentially missing X_2, the scenario is MNAR. The corresponding adjustment set remains $X_{adj} = \{X_1, X_2\}$.

Identification of $p(\bar{R} = \vec{1}|X_{adj})$:
The propensity score is still identified by

$$\begin{aligned} p(\bar{R} = \vec{1}|X_{adj}) &= p(\bar{R} = \vec{1}|X_1, X_2) \\ &= p(\bar{R}_2 = 1|X_1)p(\bar{R}_3 = 1|X_1, X_2) \\ &= p(\bar{R}_2 = 1|X_1)p(\bar{R}_3 = 1|X_1, X_2, \bar{R}_2 = 1) \end{aligned} \tag{A.6}$$

where we used the conditional independence $\bar{R}_3 \perp\!\!\!\perp \bar{R}_2 | X_1, X_2$ to make the expression a function of the observed data. The problem with using the propensity score for the AFAIS estimator is that it needs to be evaluated for all X that are sampled from $p(X)$ and not only for fully observed samples of X (as is the case in regular IPW). In this example, it is not possible to evaluate $p(\bar{R} = \vec{1} | X_{adj})$ for data points where $\bar{R}_2 = 0$. This means AFAIS cannot be used. One could impute the missing values using the MI procedure. This can, however, be very difficult due to complex identification and modelling. Furthermore, the identification of the AFA propensity score might also be very difficult or potentially impossible in these scenarios.

A.7 Derivation of the modified active feature acquisition importance sampling (AFAIS)

The derivation of the modified AFAIS estimator that combines importance sampling ideas from the black-box missing data view and the off-environment bandit view, and therefore covers the AFAIS-IPW spectrum, closely follows the derivation done in Appendix A.5. Again, the target distribution is

$$p(C, \hat{R}, X_{adj} | do(\bar{R} = \vec{1})) = p(C | X_{adj}, \hat{R}) p(\hat{R} | X_{adj}, \bar{R} = \vec{1}) p(X_{adj}) \tag{A.7}$$

Now, we propose the following behavior distribution:

$$p(C, \hat{R}, X_{adj} | \bar{R}_s = \vec{1}) = p(C | X_{adj}, \hat{R}) p(\hat{R} | X_{adj}, \bar{R}_s = \vec{1}) p(X_{adj} | \bar{R}_s = \vec{1}) \tag{A.8}$$

This means we sample from data points where some subset s of the features is fully observed ($\bar{R}_s = 1$). Sampling from Eq. (A.8) is equivalent to samping from $p(C, \hat{R}, X_{adj}, X_{\text{-}adj}, \bar{R}) = p(C, \hat{R}, X, \bar{R}) = p(C | X, \hat{R}) p(\hat{R} | X, \bar{R}) p(X, \bar{R} | \bar{R}_s = \vec{1})$ which one can do as shown in Appendix A.3.

Based on the introduced target and behavior distribution, we now derive the modified AFAIS estimator:

$$\begin{aligned}
J_{do(\bar{R}=\vec{1})} &= \mathbb{E}\left[C \middle| do(\bar{R} = \vec{1}) \right] = \sum_C C p(C | do(\bar{R} = \vec{1})) \\
&= \sum_{C, X_{adj}, \hat{R}} C p(C | X_{adj}, \hat{R}) p(\hat{R} | X_{adj}, \bar{R} = \vec{1}) p(X_{adj}) \\
&= \sum_{C, X_{adj}, \hat{R}} C p(C | X_{adj}, \hat{R}) p(\hat{R} | X_{adj}, \bar{R}_s = \vec{1}) p(X_{adj} | \bar{R}_s = \vec{1}) \cdot \frac{p(\hat{R} | X_{adj}, \bar{R} = \vec{1})}{p(\hat{R} | X_{adj}, \bar{R}_s = \vec{1})} \frac{p(X_{adj})}{p(X_{adj} | \bar{R}_s = \vec{1})} \\
&\overset{(*)}{=} \sum_{C, X_{adj}, \hat{R}} C p(C | X_{adj}, \hat{R}) p(\hat{R} | X_{adj}, \bar{R}_s = \vec{1}) p(X_{adj} | \bar{R}_s = \vec{1}) \cdot \frac{p(\bar{R} = \vec{1} | X_{adj}, \hat{R}, \bar{R}_s = \vec{1}) p(\bar{R}_s = \vec{1})}{p(\bar{R} = \vec{1} | X_{adj})} \\
&= \mathbb{E}_{C, \hat{R}, X_{adj}, \bar{R}} \left[C \frac{p(\bar{R} = \vec{1} | X_{adj}, \hat{R}, \bar{R}_s = \vec{1}) \mathbb{I}(\bar{R}_s = \vec{1})}{p(\bar{R} = \vec{1} | X_{adj})} \right] \equiv J_{AFAIS}(s)
\end{aligned} \tag{A.9}$$

where the equivalence $(*)$ holds due to Bayes rule:
The first fraction is

$$\begin{aligned}
\frac{p(\hat{R} | X_{adj}, \bar{R} = \vec{1})}{p(\hat{R} | X_{adj}, \bar{R}_s = \vec{1})} &= p(\hat{R} | X_{adj}, \bar{R} = \vec{1}) \frac{p(\bar{R} = \vec{1} | X_{adj}, \hat{R}, \bar{R}_s = \vec{1})}{p(\hat{R} | X_{adj}, \bar{R}_s = \vec{1}, \bar{R} = \vec{1}) p(\bar{R} = \vec{1} | X_{adj}, \bar{R}_s = \vec{1})} \\
&= \frac{p(\hat{R} | X_{adj}, \bar{R} = \vec{1})}{p(\hat{R} | X_{adj}, \bar{R} = \vec{1})} \frac{p(\bar{R} = \vec{1} | X_{adj}, \hat{R}, \bar{R}_s = \vec{1})}{p(\bar{R} = \vec{1} | X_{adj}, \bar{R}_s = \vec{1})} \\
&= \frac{p(\bar{R} = \vec{1} | X_{adj}, \hat{R}, \bar{R}_s = \vec{1})}{p(\bar{R} = \vec{1} | X_{adj}, \bar{R}_s = \vec{1})}
\end{aligned} \tag{A.10}$$

The second fraction is:

$$\frac{p(X_{adj})}{p(X_{adj} | \bar{R}_s = \vec{1})} = \frac{p(\bar{R}_s = \vec{1})}{p(\bar{R}_s = \vec{1} | X_{adj})} \tag{A.11}$$

Combining both Eqs.(A.10) and (A.11) gives for the importance ratio:

$$\frac{p(\hat{R}|X_{adj}, \bar{R} = \vec{1})}{p(\hat{R}|X_{adj}, \bar{R}_s = \vec{1})} \frac{p(X_{adj})}{p(X_{adj}|\bar{R}_s = \vec{1})} = \frac{p(\bar{R} = \vec{1}|X_{adj}, \hat{R}, \bar{R}_s = \vec{1})}{p(\bar{R} = \vec{1}|X_{adj}, \bar{R}_s = \vec{1})} \frac{p(\bar{R}_s = \vec{1})}{p(\bar{R}_s = \vec{1}|X_{adj})}$$
$$= \frac{p(\bar{R} = \vec{1}|X_{adj}, \hat{R}, \bar{R}_s = \vec{1}) p(\bar{R}_s = \vec{1})}{p(\bar{R} = \vec{1}|X_{adj})} \tag{A.12}$$

A.8 Identification example for the modified AFAIS

Here we show that the modified AFAIS estimator with $s = \{2\}$ is identified for the MNAR scenario from Appendix A.6 and that it is possible to evaluate the AFAIS estimator in this case. We follow the same example and see, first of all, that the propensity score now only needs to be evaluated on the datapoints where $\bar{R}_2 = 1$ which is possible. Secondly we show identification of the AFA propensity score:

Identification of $p(\bar{R} = \vec{1}|X_{adj}, \hat{R}, \bar{R}_s = \vec{1}) = p(\bar{R} = \vec{1}|X_1, X_2, \hat{R}, \bar{R}_2 = 1)$:
The identification of the AFA propensity score $p(\bar{R} = \vec{1}|X_1, X_2, \hat{R}, \bar{R}_2 = 1)$ is straight-forward, as the conditioning $\bar{R}_2 = 1$ makes this a function of observed data only.

A.9 Other views on AFAPE

Off-policy policy evaluation (OPE) for AFA

Instead of viewing observational missingness from a set view, one can also view both observational and AFA missingness jointly from a sequence view (Figure A.3). This OPE view evaluates the performance of a target policy, the AFA policy, from data collected from a behavior policy which in this case corresponds to the observational missingness (e.g. the policy the doctors had in the hospital when deciding on what features to acquire). This view differs fundamentally from the AFA graph and therefore the approach taken in this paper. The first drawback of the OPE view is that it requires additional knowledge about the order of acquisitions $\bar{U} = (\bar{U}^{(1)}, ..., \bar{U}^{(N)})$ in the retrospective dataset. Note that this drawback could be avoided, by looking at the problem as an off-policy bandit problem (instead of off-policy RL) by focusing on R instead of U, similar to the off-environment bandit view taken in this paper. For clarity, we denote the AFA missingness RL policy as $\pi_{AFA} \equiv \pi(\hat{U}^{(k)}|\hat{X}^{(k)}, \hat{R}^{(k)})$ and the observational missingness RL policy as $\pi_{Obs} \equiv \pi(\bar{U}^{(k)}|X, \bar{R}^{(k)})$. Note that $\bar{U}$ and $\hat{U}$ now describe the same variable, but the former is the observational distribution that acquired the data, while the latter is the interventional distribution that we would like to apply when deploying the AFA agent. A fundamental difference between the two distributions is, however, that $\pi_{obs} \notin \mathcal{F}_{MAR}$ in general, which implies arrows $X \to \bar{U}^{(k')}\ \forall k'$ (depicted as blue arrows in Figure A.3) that are not mediated by $\bar{X}^{(k')}$.

The evaluation of J under any AFA policy π_{AFA} under the sequence view (and Assumption 1) corresponds to:

$$J_{do(\bar{R}=\vec{1})} = \mathbb{E}[C|do(\bar{R} = \vec{1})] = \mathbb{E}[C|do(\bar{U} \sim \pi_{AFA})] \equiv J_{do(\bar{U}\sim\pi_{AFA})} \tag{A.13}$$

This reads as: *the expected cost had, instead of the observational RL policy, the AFA RL policy been applied*. This shows that the AFAPE problem can be formulated using the AFA graph (with black-box missing data and off-environment bandit views) or using this alternative OPE view. As the AFA policy does not depend on X, but only $\hat{X}$, the arrows $X \to \bar{U}^{(k')}\ \forall k'$ are eliminated under the intervention. Thus, in this setting, X is also a confounder due to for example the backdoor path $\bar{U}^{(k')} \leftarrow X \to Y \to \tilde{C}$. Therefore, unbiased evaluation requires adjustment of X which is complex due to potentially missing X values. However, adjustment is not needed under the assumption $\pi_{Obs} \in \mathcal{F}_{MAR}$, which means that the observational data was potentially also acquired by an *observational* "AFA" agent (e.g. the physician in the hospital).

A drawback of this approach is that every datapoint in the retrospective dataset is now represented by the distribution $(X, Y, \bar{U})$. This is in strong contrast to the viewpoint taken in this paper based on the AFA graph. In the AFA graph approach, one can still apply online RL and let the agent

explore how different feature combinations affect the cost, with only the restriction that acquisition actions for unavailable features are blocked. This off-policy RL setting, however, allows no active exploration as all data is retrospective with fixed trajectories of actions. It can therefore also be more accurately described as offline RL (Levine et al., 2020) because the behavior policy cannot be changed. This is a strong limitation of this viewpoint and leads to extremely low data efficiency. Furthermore, there are additional requirements for unbiased estimation of the expected cost. In the off-environment setting, the only requirement was that the propensity score is non-zero for all X (positivity assumption). In this setting, the respective positivity assumption is that the probability of all trajectories must be nonzero to be able to evaluate all possible policies, which is a much stronger assumption.

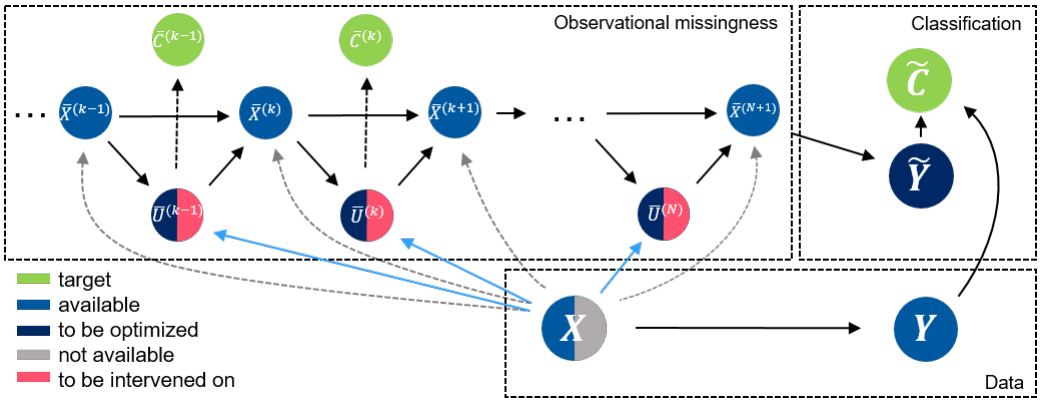

Figure A.3: AFA setting from a sequence view on AFA and observational missingness. The graph describes the observational data given by the observational missingness policy $\pi_{Obs} = \pi(\bar{U}^{(k)}|X, \bar{R}^{(k)})$. The run-time application of AFA agents then corresponds to the intervention $do(\bar{U} \sim \pi_{AFA})$ in which arrows $X \to \bar{U}$ are eliminated. This necessitates adjustment of the confounder X.

Blocking acquisition actions

The last common approach used to handling missing data in AFA settings is to block a feature acquisition action if the corresponding feature is missing (Janisch et al., 2020; Yoon et al., 2018). Yoon et al. (2018) additionally supplied the AFA agent with the missingness indicators in the retrospective dataset, making use of "informative missingness". This approach, however, corresponds to evaluating $J_{\bar{R}} \equiv \mathbb{E}[C]$ and not $J_{do(\bar{R}=\vec{1})}$ and thus is a biased estimator in the setting we consider (due to Assumption 1). Under Assumption 1, there are no "informative missingness" indicators at run-time that could be given to the agent. This method is unbiased only under the alternative assumption that the exact same restrictions for the agent at train/evaluation time apply also at run-time.

A.10 Experiment details

Synthetic data experiment

Data and costs
We evaluated and compared the described methods on a synthetic dataset with $X \in \mathbb{R}^4$ and a binary label Y, simulated using Scikit-learn's `make_classification` function (Pedregosa et al., 2011). We assumed one free feature and generated random acquisition costs c_{acq} (from a uniform distribution between 0.5 and 1) for other features. For the misclassification cost function, we used $c_{mc}(y_l, y_m) = 10$ if $l \neq m$.

Synthetic missingness mechanism
We inserted artificial MNAR missingness similar to the graph in Figure A.2B) in Appendix A.6. We assumed *i)* X_1 (fully observed) causes missingness in X_2, X_3 and X_4; *ii)* X_3 causes missingness in X_2 and X_4. We chose $s = \{3\}$ as the conditioning set when using the modified AFAIS estimator to

allow identifiability. All missingness indicators follow the logistic model (e.g. $p(\bar{R}_2 = 1|X_1, X_3) = \sigma(-4X_1 + 5X_2 + 1)$). This leads to a complete-case rate of 25%: $p(\bar{R} = \vec{1}) \approx 0.25$.

Training
We performed identification as described in Appendices A.6 and A.8 assuming the ground truth m-graph is given. We used fully-connected neural networks to train the individual terms in the propensity score (2 layers with 32 neurons each) and the AFA propensity score (3 layers with 32 neurons each). Propensity score training was done using a separate datasplit (40'000 datapoints) from the evaluation datasplit ($n_D = 100$'000 datapoints). We use an Impute-then-regress classifier (Le Morvan et al., 2021) with unconditional mean imputation and a random forest classifier for the classification task and train it on fully available and randomly subsampled data (where $p(\hat{R}_i = 1) = 0.5$). We considered the average cost when running the AFA agent on the ground truth dataset (without missingness) $J_{do(\bar{R}=\vec{1})}$ over all 100'000 datapoints as the ground truth expected cost when evaluating convergence to the correct estimate.

REAL-WORLD DATA EXPERIMENT

We tested the described estimators on real-world datasets with synthetic missingness. Here, we allowed "superfeatures" (i.e. features that comprise multiple subfeatures which are usually acquired / missing jointly and have a single cost). We further excluded CATE in these experiments as it does not allow superfeatures.

HELOC dataset
First, we tested on the HELOC (Home Equity Line of Credit) dataset which is part of the FICO explainable machine learning (xML) challenge (FICO, 2018). The dataset contains credit applications made by homeowners. The ML task was to predict based on information at the application, whether an applicant is able to repay their loan within two years. The dataset contains 9 superfeatures. We assumed a misclassification cost of 15. We induced MCAR missingness leading to $\approx 2\%$ complete cases. Note that IPW and complete case analysis coincide in this case.

Retinopathy dataset
Next, we tested on the Retinopathy dataset (Antal & Hajdu, 2014). The ML task was to predict diabetic retinopathy based on extracted image features. We considered image quality assessment and pre-screening as free features and manually extracted image features by an expert as costly. The dataset contains 7 superfeatures. We assumed a misclassification cost of 12. We induced MAR missingness, whereby other features are missing based on the pre-screening. This results in $\approx 27\%$ complete cases.

Adult income dataset
We performed evaluation on the Income dataset from the UCI data repository (Newman et al., 1998). The ML task was to predict whether a person has an income over 50'000$. We considered age as a free feature and other private information (such as education, work experience, etc.) as costly features. The dataset contains 10 superfeatures. We assumed a misclassification cost of 30. We induced MNAR missingness, whereby age and gender cause other features to be missing, but gender might be missing itself. This results in $\approx 20\%$ complete cases. We used the modified AFAIS to enable evaluation.

Training
We considered the same training as for simulation experiment except for the datasplit. We use a 50/50 datasplit where we use 50% for training the agents and propensity scores and 50% to evaluate the estimators.

A.11 EXPERIMENT RESULTS

Synthetic data experiment results are shown in the main text in Figures 5 and 6. The expected costs estimates for all real-world data experiments are shown in Figure A.4. The plot exemplifies the difference between the J_{AFAIS} and J_{IPW} estimators. While both estimators show good overlap with the ground truth, the AFAIS estimator is more data efficient. This results in much tighter boxplots for AFAIS compared to IPW. This benefit is large for agents that do not acquire much data (i.e. the random 10% agent), but diminishes as the agent becomes more "data hungry". It is thus much easier and less risk-prone to deploy an AFA agent evaluated using AFAIS. Note further, that the advantage

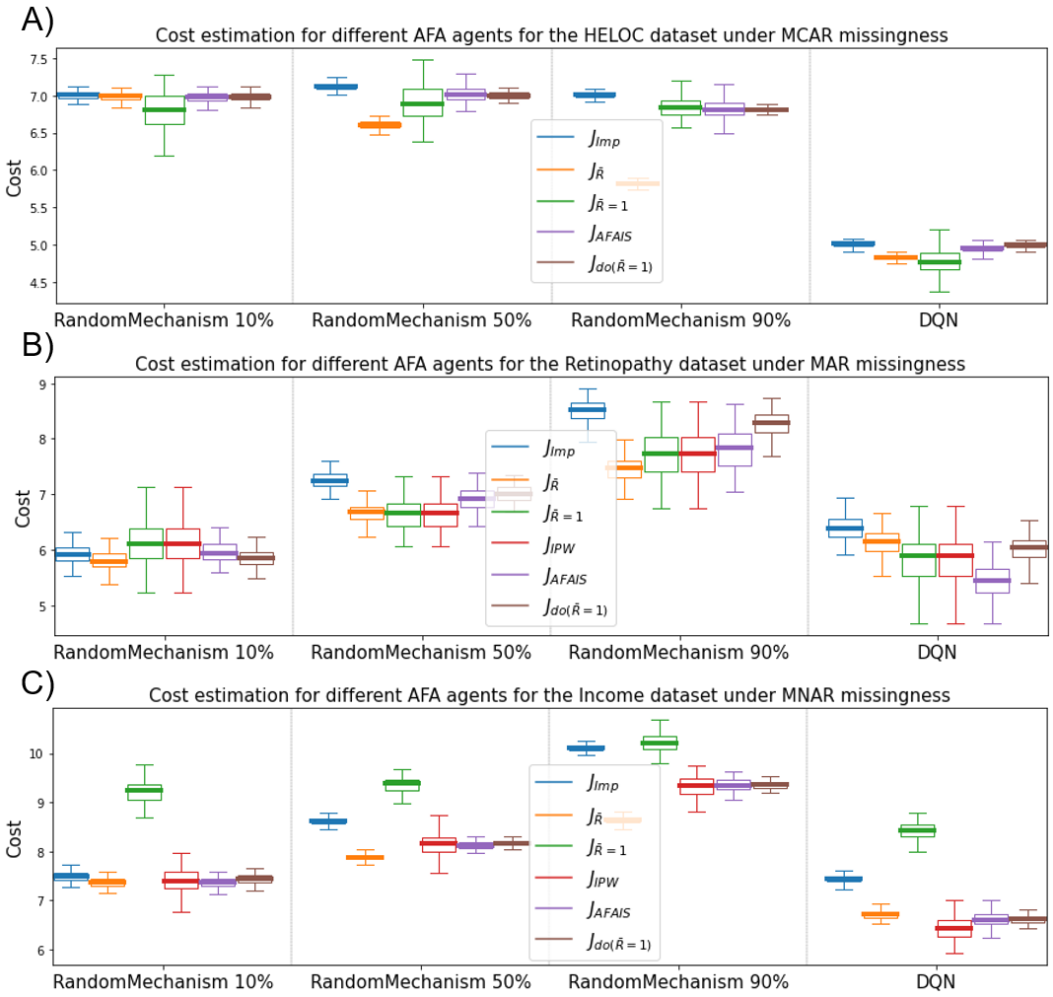

Figure A.4: Experiment results for the A) HELOC, B) Retinopathy, and C) Income datasets. Boxplots shows variability of the estimate when bootstrapping 100 times with replacement n_D datapoints from a total evaluation dataset of size n_D (n_D = 17'781 for HELOC, $n_D = 575$ for Retinopathy and n_D = 16'281 for Income dataset).

of the AFAIS estimator might, however, decrease for low sample sizes and complex agents as this increases the risk of inaccurate fitting of the AFA propensity score. This might explain the slight deviation of J_{AFAIS} from $J_{do(\bar{R}=\vec{1})}$ for the DQN agent in the Retinopathy dataset which is the smallest tested dataset with only $n_D = 575$ datapoints.

An additional benefit of IS-based estimators is that one can further use the estimator to not only estimate the counterfactual cost, but also other quantities of interest, such as counterfactual feature acquisitions. Figure A.5 shows the acquisitions of the DQN agent for the Income dataset experiment. While the sampling is done on a dataset with missingness and blocking of feature acquisitions, the AFAIS estimator (middle plot) can be used to estimate the counterfactual acquisitions that would have been made had the dataset been fully observed (right plot). This is an important insight e.g. for deployment in hospitals which need to manage their resources.

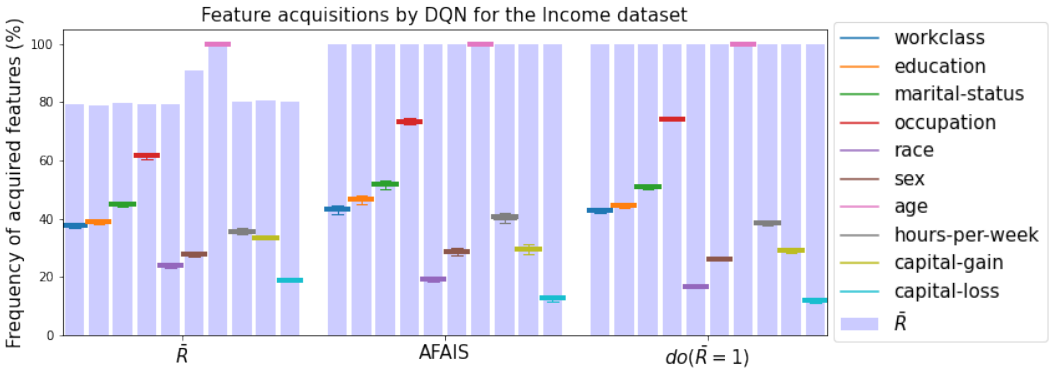

Figure A.5: Acquisitions of the DQN agent for the Income dataset. $\bar{R}$ corresponds to the acquisitions that were performed by the agent on the retrospective dataset. The acquisitions that were made by the agent when running it on the ground truth dataset (without missingness) are shown on the right ($do(\bar{R} = \vec{1})$). The middle shows the estimate by the AFAIS estimator.

A.12 GLOSSARY OF TERMS AND SYMBOLS

Term	Description
run-time	Time when the algorithm is applied prospectively (e.g. in the hospital)
observational missingness	Missingness in the retrospective dataset used to train/test the AI systems
AFA missingness	Missingness that is created at run-time when the AFA agent decides on feature acquisitions
AFAPE	Active feature acquisition performance evaluation: The problem of estimating how an AFA agent and classifier would perform at run-time (under no observational missingness) from retrospective data with observational missingness
AFA graph	Causal graph describing the AFA problem
OPE	Off-policy policy evaluation
MCAR	Missing completely at random missingness scenario
MAR	Missing at random missingness scenario
MNAR	Missing not at random missingness scenario
IPW	Inverse probability weighting
MI	Multiple imputation
AFAIS	Active feature acquisition importance sampling: novel estimator for the AFAPE problem based on an off-environment bandit view
IS	Importance sampling
sequence view	View on missingness with a focus on the sequence of feature acquisitions
set view	The traditional view on missingness with a focus on which set of features was acquired
propensity score	$p(\bar{R} = \vec{1}\|X)$ (weighting term in the IPW and AFAIS estimators)
AFA propensity score	$p(\bar{R} = \vec{1}\|X, \hat{R}, \bar{R}_s = \vec{1})$ (weighting term in the AFAIS estimators)

Symbol	Description
X	Ground truth value of features (potentially unknown)
Y	Ground truth value of the label
$\bar{R}$	Missingness indicator (observational missingness)
$\bar{X}$	Observed proxy for X (observational missingness)
$\bar{U}^{(k)}$	Feature acquisition action at acquisition step k (observational missingness)
$\hat{R}$	Missingness indicator (AFA missingness)
$\hat{X}$	Observed proxy for X (AFA missingness)
$\hat{U}^{(k)}$	Feature acquisition action at acquisition step k (AFA missingness)
π	Missingness RL policy, i.e. probability of taking a certain feature acquisition action

Symbol	Description
$\mathcal{F}_{MAR}$	Functional space for missingness policies that restricts decisions of acquiring new features to be based only on previously acquired features
c_{acq}	Vector of feature acquisition costs: element i is the cost of acquiring feature X_i
$c_{mc}(y_l, y_m)$	Misclassification cost of predicting a label of class y_l as belonging to class y_m.
$\tilde{Y} = f(\hat{R}, \hat{X})$	Predicted value for label Y using classifier f from features that have been acquired by the AFA agent
$C = \hat{C} + \tilde{C}$	Random variable describing the total cost consisting of the accumulated feature acquisition $\hat{C}$ and the misclassification cost $\tilde{C}$
$J = \mathbb{E}\left[C\right]$	Expected cost of applying an AFA missingness RL policy π and classifier f on a dataset with observational missingness
$J_{do(\bar{R}=\vec{1})} = \mathbb{E}\left[C \vert do(\bar{R} = \vec{1})\right]$	Expected cost of applying an AFA missingness RL policy π and classifier f at run-time (without observational missingness)
$h(X, Y) := \mathbb{E}[C \vert X, Y]$	Expected cost given all features and the label
J_{AFAIS}	Active feature acquisition importance sampling (AFAIS) estimator for $J_{do(\bar{R}=\vec{1})}$
J_{IPW}	Inverse probability weighting (IPW) estimator for $J_{do(\bar{R}=\vec{1})}$
J_{Imp}	Imputation-based estimator for $J_{do(\bar{R}=\vec{1})}$
$J_{\bar{R}}$	Estimator for $J_{do(\bar{R}=\vec{1})}$ based on the blocking of acquisition actions for not available features
$J_{\bar{R}=\vec{1}}$	Estimator for $J_{do(\bar{R}=\vec{1})}$ that only considers those data points where all features are available
$s \subseteq \{0, \ldots, d_x\}$	Hyperparameter of the modified AFAIS estimator that denotes which subset of the features must be complete (i.e. $\bar{R}_s = \vec{1}$) in order for a datapoint to be used in the evaluation. A smaller set s implies higher data efficiency (less restrictions on complete information), while a larger set s increases the possibility of identification and evaluation for a particular missingness scenario.