Generalization Error Bounds for Learning under Censored Feedback

Anonymous authors
Paper under double-blind review

Abstract

Generalization error bounds from learning theory provide statistical guarantees on how well an algorithm will perform on previously unseen data. In this paper, we characterize the impacts of data non-IIDness due to censored feedback (a.k.a. selective labeling bias) on such bounds. We first derive an extension of the well-known Dvoretzky-Kiefer-Wolfowitz (DKW) inequality, which characterizes the gap between empirical and theoretical CDFs given *IID* data, to problems with *non-IID data due to censored feedback*. We then use this CDF error bound to provide a bound on the generalization error guarantees of a classifier trained on such non-IID data. We show that existing generalization error bounds (which do not account for censored feedback) fail to correctly capture the model's generalization guarantees, verifying the need for our bounds. We further analyze the effectiveness of (pure and bounded) exploration techniques, proposed by recent literature as a way to alleviate censored feedback, on improving our error bounds. Together, our findings illustrate how a decision maker should account for the trade-off between strengthening the generalization guarantees of an algorithm and the costs incurred in data collection when future data availability is limited by censored feedback.

1 Introduction

Generalization error bounds are a fundamental concept in machine learning, which provide (statistical) guarantees on how a machine learning algorithm trained on some given dataset will perform on new, unseen data. However, many implicit or explicit assumptions about training data are often made when training ML models and deriving theoretical guarantees for their performance. These assumptions include access to independent and identically distributed (IID) training data, the availability of correct labels, and static underlying data distributions (Bartlett & Mendelson, 2002; Bousquet & Elisseeff, 2002; Cortes et al., 2019; 2020). Some studies in this area, e.g. Cheng et al. (2018); Kuznetsov & Mohri (2017); Mohri & Rostamizadeh (2007; 2008), have provided bounds when these assumptions are removed. In this paper, we are similarly interested in the impact of non-IID training data, specifically due to *censored feedback*, on the learned algorithm's generalization error guarantees.

Censored feedback, also known as selective labeling bias, arises in many applications wherein human or algorithmic decision-makers set certain thresholds or criteria for favorably classifying individuals, and subsequently only observe the true label of individuals who pass these requirements. For example, schools may require a minimum GPA or standardized exam score for admission; yet, graduation rates are only observed for admitted students. Financial institutions may set limits on the minimum credit score required for loan approval; yet, loan return rates are only observed for approved applicants. In these types of classification tasks, the algorithm's training dataset grows over time (as students are admitted, loans are granted); however, the new data is selected in a non-IID manner from the underlying domain, due to the unobservability of the true label of rejected data. This type of bias also arises when determining recidivism in courts, evaluating the effectiveness of medical treatments, flagging fraudulent online credit card transactions, etc. Despite this ubiquity, to the best of our knowledge, generalization error bounds given non-IID training data due to censored feedback remain unexplored. We close this gap by providing such bounds in this work, show the need for them, and formally establish the extent to which censored feedback hinders generalization.

One of the commonly proposed methods to alleviate the impacts of censored feedback is to *explore* the data domain, and admit (some of) the data points that would otherwise be rejected, with the goal of expanding the training data. Existing approaches to exploration can be categorized into *pure exploration* (Bechavod et al., 2019; Kazerouni et al., 2020; Kilbertus et al., 2020; Nie et al., 2018), where any individual in the exploration range may be admitted (with some probability ϵ), and *bounded exploration* (Balcan et al., 2007; Lee et al., 2023; Wei, 2021; Yang et al., 2022), in which the exploration range is further limited based on cost or informativeness of the new samples. The additional data samples collected through (pure or bounded) exploration may not only help improve the accuracy of the learned model when evaluated on a given test data (as shown by these prior works), but may also help tighten the generalization error guarantees of the learned model; we formalize the latter improvement, and show how the frequency and range of exploration can be adjusted accordingly.

We note that censored feedback may or may not be avoidable depending on the application (given, e.g., the costs or legal implications of exploration). We therefore present generalization error bounds both with and without exploration, establishing the extent to which the decision maker should be concerned about censored feedback's impact on the learned model's guarantees, and how well they might be able to alleviate it if exploration is feasible.

Our approach. We characterize the generalization error bounds as a function of the gap between the empirically estimated cumulative distribution function (CDF) obtained from the training data, and the ground truth underlying distribution of data. At the core of our approach is noting that although censored feedback leads to training data being sampled in a non-IID fashion from the true underlying distribution, this non-IID data can be split into IID subdomains. Existing error bounds for IID data, notably the Dvoretzky-Kiefer-Wolfowitz (DKW) inequality (Dvoretzky et al., 1956; Massart, 1990), can provide bounds on the deviation of the empirical and theoretical subdomain CDFs, as a function of the number of available data samples in each subdomain. The challenge, however, lies in reassembling such subdomain bounds into an error bound on the full domain CDFs. Specifically, this will require us to shift and/or scale the subdomain CDFs, with shifting and scaling factors that are themselves empirically estimated from the underlying data, and can be potentially re-estimated as more data is collected. Our analysis identifies these factors, and highlights the impacts of each on the error bounds.

Summary of findings and contributions:

1. We generalize the well-known Dvoretzky-Kiefer-Wolfowitz (DKW) inequality, which characterizes the gap between empirical and theoretical CDFs given *IID* data, to problems with *non-IID data due to censored feedback* without exploration (Theorem 2) and with exploration (Theorem 3), and formally show the extent to which censored feedback hinders generalization.
2. We characterize the change in these error bounds as a function of the severity of censored feedback (Proposition 1) and the exploration frequency (Proposition 2). We further show (Section 3.3) that a minimum level of exploration is needed to tighten the error bound.
3. We derive a generalization error bound (Theorem 4) for a classification model learned in the presence of censored feedback using the CDF error bounds in Theorems 2 and 3.
4. We numerically illustrate our findings (Section 5). We show that existing generalization error bounds (which do not account for censored feedback) fail to correctly capture the generalization error guarantees of the learned models. We also illustrate how a decision maker should account for the trade-off between strengthening the generalization guarantees of an algorithm and the costs incurred in data collection for reaching enhanced learning guarantees.

Related works. Although existing literature has studied generalization error bounds for learning from non-IID data, non-IIDness raised by censored feedback has been overlooked. We provide a detailed review of related work in Appendix A. Here, we discuss works most closely related to ours.

First, our work is closely related to generalization theory in the PAC learning framework in non-IID settings, including (Mohri & Rostamizadeh, 2007; 2008; Yu, 1994) and (Kuznetsov & Mohri, 2017); these works consider dependent samples generated through a stationary, and non-stationary β-mixing sequence, respectively, where the dependence between samples weakens over time. To address the vanishing dependence issue, these works consider building blocks within which the data can be viewed as IID. The study of Yu (1994) is based on the VC-dimension, while Mohri & Rostamizadeh (2008) and Mohri & Rostamizadeh (2007) focus on the Rademacher complexity and algorithm stability, respectively. Our work is similar in that we also consider building IID blocks to circumvent data non-IIDness. However, we differ in our reassembly method, in the source of non-IID data, and in our study of the impacts of exploration.

Our work is also closely related to partitioned active learning, including Cortes et al. (2019; 2020); Lee et al. (2023); Zheng et al. (2019). Cortes et al. (2019) partition the entire domain to find the best hypothesis for each subdomain, and a PAC-style generalization bound is derived compared to the best hypothesis over the entire domain. This idea is further extended to adaptive partitioning in Cortes et al. (2020). In Lee et al. (2023), the domain is partitioned into a fixed number of subdomains, and the most uncertain subdomain is explored to improve the mean-squared error. The work of Zheng et al. (2019) considers a special data non-IIDness where the data-generating process depends on the task property, partitions the domain according to the task types, and analyzes each subdomain separately. Our work is similar to these studies in that we also consider (active) exploration techniques, and partition the data domain to build IID blocks. However, we differ in problem setup and analysis approach, and in accounting for the cost of exploration when we consider bounded exploration techniques.

Lastly, the technique of identifying IID-blocks within non-IID datasets has also been used in other contexts to address the challenge of generalization guarantees given non-IID data. For instance, Wang et al. (2023) investigate generalization performance with covariate shift and spatial autocorrelation in geostatistical learning. They address the non-IIDness issue by removing samples from the buffer zone to construct spatially independent folds. Similarly, Tang et al. (2021) study generalization performance within the Federated Learning paradigm with non-IID data. They employ clustering techniques to partition clients into distinct clusters based on statistical characteristics, thus treating samples from clients within each cluster as IID and analyzing each cluster separately. We similarly explore generalization performance with non-IID data samples and employ the technique of identifying IID subdomains/blocks. However, we differ in the reason for the occurrence of non-IIDness, the setup of the problem, and our analytical approaches.

2 Problem Setting

Consider a decision maker (equivalently, the algorithm), and new agents (equivalently, data points) arriving sequentially. The algorithm is a binary classifier, used to make positive/negative decisions (e.g., accept/reject) on each new data point. We use a bank granting loans as a running example.

The agents. Each agent/data point has a feature x and a true label y. The feature $x \in \mathcal{X} \subseteq \mathbb{R}$ is the characteristic used for decision-making (e.g., a credit score). The true label $y \in \mathcal{Y} = \{0, 1\}$ reflects a qualification state, with $y = 1$ meaning the data point is qualified to receive a favorable decision (e.g., the applicant will return a loan if granted). We will use X and Y to denote the corresponding random variables, and x and y to denote realizations. Denote the proportion of qualified (unqualified) samples in the population by p_1 (p_0).

The algorithm. The decision maker begins with a fixed initial/historical training dataset containing n_y IID samples [1] $\{x^y_1, \dots, x^y_{n_y}\}$ for each label y (e.g., data on past loan repayments with n_0 IID samples of individuals with credit scores $\{x^0_1, \dots, x^0_{n_0}\}$ who defaulted on their loans, and n_1 IID samples of individuals with credit scores $\{x^1_1, \dots, x^1_{n_1}\}$ who paid off their loans on time). Based on these, the decision maker selects a threshold-based binary classifier $f_\theta(x) : \mathcal{X} \to \{0, 1\}$ to decide whether to admit or reject incoming agents (equivalently, assign labels 1 or 0). Specifically, $f_\theta(x) = \mathbb{1}(x \geq \theta)$, with θ denoting the decision threshold (e.g., θ could be the minimum credit score to be approved for a loan).[2]

---

[1] That is, we assume that any non-IIDness is introduced due to censored feedback impacting subsequent data collection. Extension to initially biased training data is also possible, but at the expense of additional notation.

[2] The single-dimensional features and threshold classifier assumptions are not too restrictive: Corbett-Davies et al. (2017, Thm 3.2) and Raab & Liu (2021) have shown that threshold classifiers can be optimal if multi-dimensional features can be appropriately converted into a one-dimensional scalar (e.g., with a neural network).

The decision threshold θ divides the data domain into two regions: the upper, *disclosed* region, where the true label of future admitted agents will become known to the decision maker, and the lower, *censored* region, where true labels are no longer observed. As new agents arrive, due to this censored feedback, additional data is only collected from the disclosed region of the data domain (e.g., we only find out if an agent repays the loan if it is granted the loan in the first place). This is what causes the non-IIDness of the (expanded) dataset: after new agents arrive, the training dataset consists of n_y historical IID samples from both censored and disclosed regions on each label y, and an additional k_y samples collected afterwards from each label y, but only from the disclosed region, making the entire $n_y + k_y$ samples a non-IID subset of the respective label y's data domain.

Formally, let $F^y(x)$ denote the theoretical (ground truth) feature distribution for label y agents. Let $\alpha^y := F^y(\theta)$ be the theoretical fraction of the agents in the censored region, and m_y be the number of the initial n_y training samples from label y agents located in the censored region. It is worth noting that $\frac{m_y}{n_y}$ can provide an empirical estimate of α^y, but the two are in general not equal. Let k_y denote the number of additional training samples on label y that have been collected, all from the disclosed region, after new agents arrive. The decision maker will now have access to $n_y + k_y$ total samples from label y agents, which are not identically distributed (m_y in the censored region, and $n_y - m_y + k_y$ in the disclosed region). Let $F^y_{n_y+k_y}(x)$ denote the empirical CDF of the feature distribution for label y agents based on these $n_y + k_y$ training data points. Our first goal is to provide an error bound, similar to the DKW inequality, of the discrepancy between F^y_{n+k} and the ground truth CDF F^y, for each label y. We will then use these to bound the generalization error guarantees of the learned model from the (non-IID) $\{n_y + k_y\}_{y\in\{0,1\}}$ data points.

Remark 1. *Note that we assume the decision maker starts with one* fixed realization *of* $\{n_y\}_{y\in\{0,1\}}$ *data points, and therefore the decision threshold* θ *and the number of initial samples in the censored region* $\{m_y\}_{y\in\{0,1\}}$ *are (non-random) fixed values. For instance, the fixed training dataset and decision threshold can be likened to a financial institution with an existing loan repayment history dataset consisting of a set number of defaults and on-time payoffs, and its initial decision threshold selected for approving future applications. However, there are two possible ways to interpret the additional samples* $\{k_y\}_{y\in\{0,1\}}$*:* a posteriori *(i.e., outcomes after collecting exactly* k_y *new samples in the disclosed region), or* a priori *(i.e., possible values once a total of* T *new agents arrive, only some of which will fall in the disclosed region). The former is a reasonable assumption if a decision maker has already collected samples under censored feedback, or alternatively, is willing to wait to collect the exact required number of samples until it can achieve a desired error bound. The latter is from the viewpoint of a decision maker contemplating potential outcomes if it waits for a total of* T *new agents to arrive. We will present our new error bound under both interpretations.*

We summarize our algorithm dynamics and main notation below:

Stage I: initial data. The decision maker starts with n_y fixed data points $\{x^y_1, \ldots, x^y_{n_y}\}$ from each label $y \in \{0, 1\}$, drawn IID from the respective true underlying distribution $F^y(x)$. Accordingly, it selects a fixed decision threshold θ. Given θ, the n_y samples from label y agents can be divided into m_y samples below θ ($m_y = |\{i : x^y_i < \theta\}|$; refered to as the censored region) and $n_y - m_y$ samples above θ (referred to as the disclosed region).

Stage II: arrival of new agents. At each time t, a new agent arrives. Its true label is $\hat{y} = y$ with probability p_y, and its feature $\hat{x}$ is drawn uniformly at random from the corresponding $F^{\hat{y}}(x)$. The agent's feature $\hat{x}$ is observed, and the agent is admitted if and only if $\hat{x} \geq \theta$. Due to censored feedback, $\hat{y}$ will only be observed if the agent is admitted. When an agent is admitted, its data is used to expand the corresponding dataset of $y = \hat{y}$ samples to $\{x^y_1, \ldots, x^y_{n_y}, x^y_{n_y+1}, \ldots, x^y_{n_y+k^t_y-1}, \hat{x}\}$.

Stage III: updating empirical distribution estimates. After T time steps (which can be fixed in advance, or denote the time at which a certain number of samples have been collected), the decision maker has access to k_y new samples of label y agents, having expanded its trainig data on label y agents to (the non-IID collection) $\{x^y_1, \ldots, x^y_{n_y}, x^y_{n_y+1}, \ldots, x^y_{n_y+k_y-1}, x^y_{n_y+k_y}\}$. Accordingly, it will find $F^y_{n_y+k_y}(x)$, the empirical CDF of the feature distribution for label y agents based on the total of $n_y + k_y$ data points on those agents.

3 Error Bounds on Cumulative Distribution Function Estimates

Recall that our first goal is to provide an error bound, similar to the DKW inequality, of the discrepancy between the empirical CDF of feature distribution $F^y_{n_y+k_y}(x)$ and the ground truth CDF $F^y(x)$, for each label y. Note that the empirical CDF is found for each label y separately based on its own data samples. Therefore, we drop the label y from our notation throughout this section for simplicity. Further, we first derive the *a posteriori* bounds for given realizations of k_y, and develop the *a priori* version of the bound accordingly in Corollary 1.

We first state the Dvoretzky-Kiefer-Wolfowitz inequality (an extension of the Vapnik–Chervonenkis (VC) inequality for real-valued data) which provides a CDF error bound given IID data.

Theorem 1 (The Dvoretzky-Kiefer-Wolfowitz (DKW) inequality (Dvoretzky et al., 1956; Massart, 1990))**.** *Let $Z_1, \ldots, Z_n$ be IID real-valued random variables with cumulative distribution function $F(z) = \mathbb{P}(Z_1 \leq z)$. Let the empirical distribution function be $F_n(z) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(Z_i \leq z)$. Then, for every n and η> 0,*

$$\mathbb{P}\left(\sup_{z\in\mathbb{R}} \Big|F(z) - F_n(z)\Big| \geq \eta\right) \leq 2\exp\left(-2n\eta^2\right) .$$

In words, the DKW inequality shows how the likelihood that the maximum discrepancy between the empirical and true CDFs exceeds a tolerance level η decreases in the number of (IID) samples n.

We now extend the DKW inequality to the case of non-IID data due to censored feedback. We do so by first splitting the data domain into blocks containing IID data, to which the DKW inequality is applicable. Specifically, although the expanded training dataset is non-IID, the decision maker has access to m IID samples in the censored region, and $n - m + k$ IID samples in the disclosed region. Let G_m and K_{n-m+k} denote the corresponding empirical feature distribution CDFs. The DKW inequality can be applied to bound the difference between these empirical CDFs and the corresponding ground truth CDFs G and K.

It remains to identify a connection between the full CDF F, and G (the censored CDF) and K (the disclosed CDF), to reach a DKW-type error bound on the full CDF estimate (see Figure 1 for an illustration). This reassembly from the bounds on the IID blocks into the full data domain is however more involved, as it requires us to consider a set of scaling and shifting factors, which are themselves empirically estimated and different from the ground truth values. We will account for these differences when deriving our generalization of the DKW inequality, as detailed in the remainder of this section. All proofs are given in the Appendix.

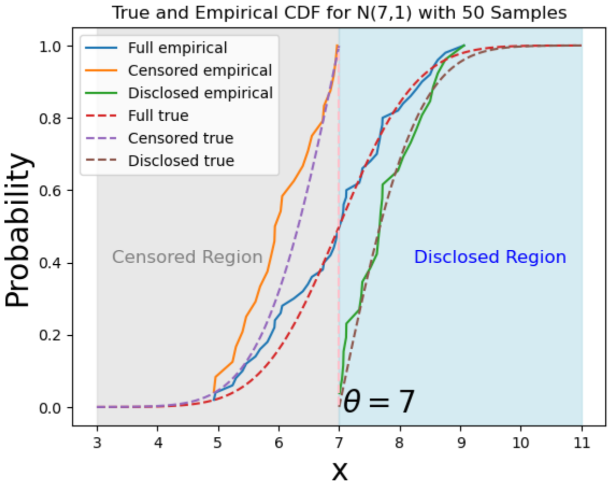

Figure 1: The empirical CDFs F_{n+k} (Full domain), G_m (Censored region), and K_{n-m+k} (Disclosed region), and the theoretical CDFs of F, G, and K. Experiments based on randomly drawn samples from Gaussian data $N(7,1)$, $\theta = 7$, $n = 50$, $m = 24$, and $k = 0$.

3.1 CDF bounds under censored feedback

We first present two lemmas that establish how the deviation of G_m and K_{n-m+k} from their corresponding theoretical values relate to the deviation of the full empirical CDF F_{n+k} from its theoretical value F.

Lemma 1 (Censored Region)**.** *Let $Z = \{X_i | X_i \leq \theta\}$ denote the m out of $n+k$ samples that are in the censored region. Let G and G_m be the theoretical and empirical CDFs of Z, respectively. Then,*

$$\sup_{x\in(-\infty,\theta)} |F(x) - F_{n+k}(x)| \leq \sup_{x\in(-\infty,\theta)} \underbrace{\left| \min\left(\alpha, \frac{m}{n}\right)(G(x) - G_m(x)) \right|}_{\textit{(scaled) censored subdomain error}} + \underbrace{\left|\alpha - \frac{m}{n}\right|}_{\textit{scaling error}} .$$

The (partial) error bound in this lemma shows the maximum difference between the true F and the empirical F_{n+k} in the censored region (i.e., for $x \in (-\infty, \theta)$) can be bounded by the maximum difference between G and G_m, modulated by the *scaling* ($\min(\alpha, \frac{m}{n})$) that is required to map from partial CDFs to full CDFs.

Specifically, to match the partial and full CDFs, we need to consider the different endpoints of the censored region's CDF and the full CDF at θ, which are $G_m(\theta) = G(\theta) = 1$, $F(\theta) = \alpha$, and $F_{n+k}(\theta) = \frac{m}{n}$, respectively. The first term in the bound above accounts for this by scaling the deviation between the true and empirical partial CDF accordingly. The second term accounts for the error in this scaling since the empirical estimate $\frac{m}{n}$ is generally not equal to the true endpoint α.

The following is a similar result in the disclosed region.

Lemma 2 (Disclosed Region)**.** *Let $Z = \{X_i | X_i \geq \theta\}$ denote the $n-m+k$ out of the $n+k$ samples in the disclosed region. Let K and K_{n-m+k} be the theoretical and empirical CDFs of Z, respectively. Then,*

$$\sup_{x\in(\theta,\infty)} |F(x) - F_{n+k}(x)| \leq \sup_{x\in(\theta,\infty)} \underbrace{\left| \min(1-\alpha, 1-\frac{m}{n})(K(x) - K_{n-m+k}(x)) \right|}_{\textit{(scaled) disclosed subdomain error}} + \underbrace{2\left|\alpha - \frac{m}{n}\right|}_{\textit{shifting and scaling errors}}$$

Similar to Lemma 1, we observe the need for a scaling factor. However, in contrast to Lemma 1, this lemma introduces an additional *shifting error*, resulting in a factor of two in the last term $|\alpha - \frac{m}{n}|$. In particular, we need to consider the different starting points of the disclosed region's CDF and full CDF at θ, which are $K_m(\theta) = K(\theta) = 0$, $F(\theta) = \alpha$, and $F_{n+k}(\theta) = \frac{m}{n}$, respectively, when mapping between the CDFs; one of the $|\alpha - \frac{m}{n}|$ captures the error of shifting the starting point of the partial CDF to match that of the full CDF.

We can now state our main theorem, which generalizes the well-known DKW inequality to problems with censored feedback.

Theorem 2. *Let $x_1, x_2, \ldots, x_n$ be fixed initial data samples, drawn IID from a distribution with CDF $F(x)$. Let θ partition the data domain into two regions, such that $\alpha = F(\theta)$, and m of the initial n samples are located to the left of θ. Assume we have collected k additional samples above the threshold θ, and let F_{n+k} denote the empirical CDF estimated from these $n+k$ (non-IID) data. Then, for every $\eta > 0$,*

$$\mathbb{P}\left[\sup_{x\in\mathbb{R}} \left|F(x) - F_{n+k}(x)\right| \geq \eta\right] \leq \underbrace{2\exp\left(\frac{-2m(\eta - |\alpha - \frac{m}{n}|)^2}{\min\left(\alpha, \frac{m}{n}\right)^2}\right)}_{\textit{censored region error (constant)}} + \underbrace{2\exp\left(\frac{-2(n-m+k)(\eta - 2|\alpha - \frac{m}{n}|)^2}{\min\left(1-\alpha, \frac{n-m}{n}\right)^2}\right)}_{\textit{disclosed region error (decreasing with additional data)}}$$

The proof proceeds by applying the DKW inequality to each subdomain, and combining the results using a union bound on the results of Lemmas 1 and 2.

The expression above shows that as the number of samples collected under censored feedback increases ($k \to \infty$), the disclosed region's error decreases exponentially (similar to the DKW bound). However, unlike the DKW bound, this error bound does not go to zero due to a constant error term from the censored region of the data domain (the first term in the error bound). This means that unless exploration strategies are adopted, we can not guarantee arbitrarily good generalization in censored feedback tasks. Finally, we note that the DKW inequality can be recovered as the special case of our Theorem 2 by letting $\theta \to -\infty$ (which makes $\alpha \approx 0, m \approx 0$).

Finally, recall from Remark 1 that, instead of considering an exact realization of k new samples in the disclosed region, a decision maker may want to know the error bound after waiting for T agents to arrive (only some of which will fall in the disclosed region). The following corollary provides an error bound that can be leveraged under this viewpoint.

Corollary 1. *Let $x_1, x_2, \ldots, x_n$ be fixed initial data samples, drawn IID from a distribution with CDF $F(x)$. Let θ partition the data domain into two regions, such that $\alpha = F(\theta)$, and m of the initial n samples are located to the left of θ. Assume we have waited for T additional samples to arrive, and let F_{n+T} denote the empirical CDF estimated accordingly. Then, for every $\eta > 0$,*

$$\mathbb{P}\left[\sup_{x\in\mathbb{R}}\left|F(x) - F_{n+T}(x)\right| \geq \eta\right] \leq \underbrace{2\exp\left(\frac{-2m(\eta - |\alpha - \frac{m}{n}|)^2}{\min\left(\alpha, \frac{m}{n}\right)^2}\right)}_{\textit{censored region error (constant)}}$$
$$+ \underbrace{\sum_{k=0}^{T} 2\binom{T}{k}(1-\alpha)^k \alpha^{T-k} \exp\left(\frac{-2(n-m+k)(\eta - 2|\alpha - \frac{m}{n}|)^2}{\min\left(1-\alpha, \frac{n-m}{n}\right)^2}\right)}_{\textit{expected disclosed region error (decreasing with wait time } T\textit{)}}$$

The proof is straightforward, and follows from writing the law of total probability for the left-hand side of the inequality by conditioning on the realization k of the samples in the disclosed region. We first note that the constant error term, as expected, is unaffected by the wait time T. The second term is the expected value of the disclosed region error from Theorem 2; it is decreasing with T as the exponential error terms decrease with k, and higher k's are more likely at higher T.

3.2 Censored feedback and exploration

A commonly proposed method to alleviate censored feedback, as noted in Section 1, is to introduce exploration in the data domain. From the perspective of the generalization error bound, exploration has the advantage of reducing the constant error term in Theorem 2, by collecting more data samples from the censored region. Formally, we consider (bounded) exploration in the *range* $x \in (LB, \theta)$, where samples in this range are admitted with an exploration *frequency* ϵ. When $LB \to -\infty$, this is a pure exploration strategy.

Now, the lowerbound LB and the decision threshold θ partition the data domain into three IID subdomains (see Figure 2 for an illustration). However, the introduction of the additional *exploration region* (LB, θ) will enlarge the CDF bounds, as it introduces new scaling and shifting errors when reassembling subdomain bounds into full domain bounds.

Specifically, of the n initial data, let l, $m - l$, and $n - m$ of them be in the censored (below LB), exploration (between LB and θ), and disclosed (above θ) regions, respectively. Let $\beta = F(LB)$ and $\alpha = F(\theta)$, with initial empirical estimates $\frac{l}{n}$ and $\frac{m}{n}$, respectively.

As new agents arrive, let k_1 and k_2 denote the additional samples collected in the exploration range and disclosed range, respectively. One main difference of this setting with that of Section 3.1 is that as additional samples are collected, the empirical estimate of α can be re-estimated. Accordingly, we present a lemma similar to Lemmas 1 and 2 for the exploration region.

Lemma 3 (Exploration Region)**.** *Let $Z = \{X_i | LB \leq X_i \leq \theta\}$ denote the $m - l + k_1$ samples out of the $n + k_1 + k_2$ samples that are in the exploration range. Let E and E_{m-l+k_1} be the theoretical and empirical CDFs of Z, respectively. Then,*

$$\sup_{x\in(LB,\theta)} |F(x) - F_{n+k_1+k_2}(x)| \leq \underbrace{\left|\beta - \frac{l}{n}\right|}_{\textit{shifting error}} + \underbrace{\left|\alpha - \beta - \frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\right|}_{\textit{re-estimated scaling error}}$$
$$+ \underbrace{\sup_{x\in(LB,\theta)}\left|\min\left(\alpha-\beta, \frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\right)(E(x) - E_{m-l+k_1}(x))\right|}_{\textit{scaled exploration subdomain error}}$$

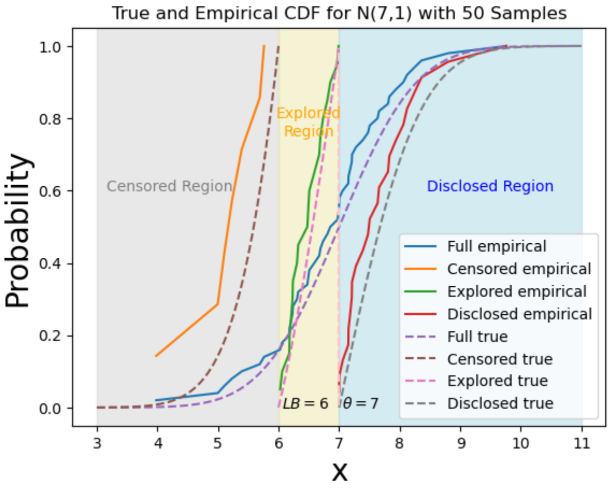

Figure 2: The empirical CDFs $F_{n+k_1+k_2}$ (Full domain), G_l (Censored region), E_{m-l+k_1} (Explored region), and K_{n-m+k_2} (Disclosed region), and the theoretical CDFs of F, G, E, and K. Experiments based on randomly drawn samples from Gaussian data $N(7, 1)$, $\theta = 7$ $LB = 6$, $n = 50$, $l = 7, m = 27$, and $k_1 = k_2 = 0$.

Observe that here, we need both scaling and shifting factors to relate the partial and full CDF bounds, as in Lemma 2, but with an evolving scaling error as more data is collected. In particular, the initial empirical estimate $\frac{m}{n}$ is updated to $\frac{l}{n} + \frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}$ after the observation of the additional k_1 and k_2 samples.

We now extend the DKW inequality when data is collected under censored feedback *and* with exploration.

Theorem 3. *Let $x_1, x_2, \ldots, x_n$ be fixed initial data samples, drawn IID from a distribution with CDF $F(x)$. Let LB and θ partition the domain into three regions, such that $\beta = F(LB)$ and $\alpha = F(\theta)$, with l and m of the initial n samples located to the left of LB and θ, respectively. Assume we have collected an additional k_1 samples between LB and θ, under an exploration probability ϵ, and an additional number of k_2 samples above θ. Let $F_{n+k_1+k_2}$ denote the empirical CDF estimated from these $n + k_1 + k_2$ non-IID samples. Then, for every $\eta > 0$,*

$$\mathbb{P}\left[\sup_{x\in\mathbb{R}}\left|F(x) - F_{n+k_1+k_2}(x)\right| \geq \eta\right] \leq 2\exp\left(\frac{-2l(\eta-|\beta-\frac{l}{n}|)^2}{\min\left(\beta,\frac{l}{n}\right)^2}\right)$$
$$+\,2\exp\left(\frac{-2(m-l+k_1)\left(\eta-|\beta-\frac{l}{n}|-\left|\alpha-\beta-\frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\right|\right)^2}{\min\left(\alpha-\beta,\frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\right)^2}\right) + 2\exp\left(\frac{-2(n-m+k_2)\left(\eta-2\left|\alpha-\frac{l}{n}-\frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\right|\right)^2}{\min\left(1-\alpha,\frac{n-l}{n}\frac{n-m+\epsilon k_2}{n-l+k_1+\epsilon k_2}\right)^2}\right).$$

Comparing this expression with Theorem 2, we first note that the last terms corresponding to the disclosed region are similar when setting $k = k_2$, with the difference being in the impact of re-estimating α.

The key difference between the two error bounds is in the censored region, in that the first term in Theorem 2 is now broken into two parts: (still) censored region $(-\infty, LB)$, and the exploration region (LB, θ). We can see that although there can still be a non-vanishing error term in the (still) censored region, as we collect more samples ($k_1 \to \infty$) in the exploration region, the error from the exploration region will decrease to zero. Further, if we adopt pure exploration ($LB \to -\infty$, which makes $\beta \approx 0, l \approx 0$), the first term will vanish as well (however, note that pure exploration may not be a feasible option if exploration is highly costly). Lastly, we note that an *a priori* version of this bound can be derived using similar techniques to that of Corollary 1.

3.3 When will exploration improve generalization guarantees?

It might seem at first sight that the new vanishing error term in the exploration range of Theorem 3 necessarily translates into a tighter error bound than that of Theorem 2 when exploration is introduced. Nonetheless, the shifting and scaling factors, as well as the introduction of an additional union bound, enlarge the CDF error bound. Therefore, in this section, we elaborate on the trade-off between these factors, and evaluate when the benefits of exploration outweigh its drawbacks in providing error bounds on the data CDF estimates.

We begin by presenting two propositions that assess the change in the bounds of Theorems 2 and 3 as a function of the severity of censored feedback (as measured by θ) and the exploration frequency ϵ.

Proposition 1. *Let $B(\theta)$ denote the error bound in Theorem 2, and assume the conditions of that theorem hold. Assume also that we can collect an additional $k = O(n)$ samples above the threshold. Then, $B(\theta)$ is increasing in θ.*

Proposition 2. *Let $B^e(LB, \theta, \epsilon)$ denote the error bound in Theorem 3, and assume the conditions of that theorem hold. Then, $B^e(LB, \theta, \epsilon)$ is decreasing in ϵ.*

In words, as intuitively expected, these propositions state that the generalization bounds worsen (i.e., are less tight) when the censored feedback region is larger, and that they can be improved (i.e., made more tight) as the frequency of exploration increases.

Numerical illustration. We also conduct a numerical experiment to illustrate the bounds derived in Theorems 2 and 3. We proceed as follows: 8000 random samples are drawn from a Gaussian distribution with mean $\mu = 7$ and standard deviation $\sigma = 3$, with an additional 40000 samples arriving subsequently, randomly sampled from across the entire data domain. We set $\eta = 0.015$, the threshold $\theta = 8$, and the lower bound $LB = 6$. We run the experiment 5 times and report the error bounds accordingly.

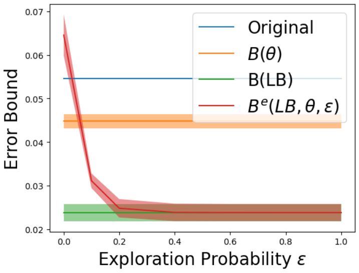

Figure 3: A minimum exploration frequency is needed to tighten the CDF error bound.

In Figure 3, the "original" (blue) line represents the DKW CDF bound of the initial samples without additional data. The "$B(\theta)$" (orange) line and "$B(LB)$" (green) line represent the CDF bound in Theorem 2 without exploration, where the decision threshold is at θ and LB, respectively. The "$B^e(LB, \theta, \epsilon)$" (red) line represents the bound in Theorem 3 with exploration probability ϵ.

From Figure 3, we first observe that the green line ($B(LB)$, which observes new samples with $x \geq LB = 6$) provides a tighter bound than the orange line ($B(\theta)$, which observes new samples with $x \geq \theta = 8$), with both providing tighter bounds than the blue line (original DKW bound, before any new samples are observed). This is because collecting more samples from the disclosed region results in a decrease in the CDF error bound, as noted by Proposition 1. Additionally, we can observe from the trajectory of the red line ($B^e(LB, \theta, \epsilon)$, which observes a fraction ϵ of new samples from (LB, θ), and all new samples above θ) that introducing exploration enlarges the CDF error bound due to the additional union bound, but it also enables the collection of more samples, leading to a decrease in the CDF error bound as ϵ increases; note that this observation aligns with Proposition 2.

Notably, we see that a minimum level of exploration probability ϵ (accepting around 10% of the samples in the exploration range) is needed to improve the CDF bounds over no exploration. Note that this may or may not be feasible for a decision maker depending on the costs of exploration (see also Section 3.4). However, if exploration is feasible, we also see that accepting around 20% of the samples in the exploration range (when the red line is close to the green line) can be sufficient to provide bounds nearly as tight as observing all samples in the exploration range.

3.4 How to choose an exploration strategy?

We close this section by discussing potential considerations in the choice of an exploration strategy in light of our findings. Specifically, a decision maker can account for a tradeoff between *the costs of exploration* and *the improvement in the generalization error bound* when choosing its exploration strategy. Recall that the exploration strategy consists of selecting an exploration lowerbound/range LB and an exploration probability

ϵ. Formally, the decision maker can solve the following optimization problem to choose these parameters:

$$\max_{\epsilon\in[0,1], LB\in[0,\theta]} \quad \big(B(\theta) - B^e(LB,\theta,\epsilon)\big) - C(LB,\theta,\epsilon) \ , \tag{1}$$

where $B(\theta)$ and $B^e(LB,\theta,\epsilon)$ denote the error bounds in Theorems 2 and 3, respectively, and $C(LB,\theta,\epsilon)$ is an exploration cost which is non-increasing in $(\theta - LB)$ (reducing the exploration range will weakly decrease the costs) and non-decreasing in ϵ (exploring more samples will weakly increase the cost). As an example, the cost function $C(LB,\theta,\epsilon)$ can be given by

$$C(LB,\theta,\epsilon) = \epsilon \int_{LB}^{\theta} e^{\frac{\theta-x}{c}} f^0(x)\mathrm{d}x. \tag{2}$$

In words, unqualified (costly) samples at x have a density $f^0(x)$, and when selected (as captured by the ϵ multiplier), they incur a cost $e^{\frac{\theta-x}{c}}$, where $c > 0$ is a constant. Notably, observe that the cost is increasing as the sample x gets further away from the threshold θ. For instance, in the bank loan example, this could capture the assumption that individuals with lower credit scores default on a larger portion of their loans.

As noted in Proposition 2, $B^e(LB,\theta,\epsilon)$ is decreasing in ϵ; coupled with any cost function $C(LB,\theta,\epsilon)$ that is (weakly) increasing in ϵ, this means that the decision maker's objective function in equation 1 captures a tradeoff between reducing generalization errors and modulating exploration costs.

The optimization problem in equation 1 can be solved (numerically) by plugging in for the error bounds from Theorems 2 and 3 and an appropriate cost function (e.g., equation 2). For instance, in the case of the numerical example of Fig. 3, under the cost function of equation 2 with $c = 5$, and fixing $LB = 6$, the decision maker should select $\epsilon = 11.75\%$.

Another potential solution for modulating exploration costs is to use multiple exploration subdomains, each characterized by an exploration range $[LB_i, LB_{i-1})$, and with a higher exploration probability ϵ_i assigned to the subdomains closer to the decision boundary (which are less likely to contain high cost samples). For instance, with the choice of b subdomains, the cost function of equation 2 would change to (the lower) cost:

$$C(\{LB_i\}_{i=1}^b, \theta, \{\epsilon_i\}_{i=1}^b) = \sum_{i=1}^{b} \epsilon_i \int_{LB_i}^{LB_{i-1}} e^{\frac{\theta-x}{c}} f^0(x)\mathrm{d}x. \tag{3}$$

It is worth noting that while this approach can reduce the costs of exploration, it will also weaken generalization guarantees when we reassemble the b exploration subdomains' bounds back into an error bound of the full domain (similar to what was observed in Fig. 3 for $b = 1$). This again highlights a tradeoff between improving learning error bounds and restricting the costs of data collection.

4 Generalization Error Bounds under Censored Feedback

In this section, we use the CDF error bounds from Section 3 to characterize the generalization error of a classification model that has been learned from data collected under censored feedback. Specifically, we will first establish a connection between the generalization error of a classifier (the quality of its learning) and the CDF error bounds on its training dataset (the quality of its data). With this relation in hand, we can then use any of the CDF error bounds from Theorems 1-3 to bound how well algorithms learned on data suffering from censored feedback (and without or with exploration) can generalize to future unseen data.

Formally, we consider a 0-1 learning loss function $\mathcal{L} : \mathcal{Y} \times \mathcal{Y} \to \{0, 1\}$. Denote $R(\theta) = \mathbb{E}_{XY}\mathcal{L}(f_\theta(X), Y)$ as the expected risk incurred by an algorithm with a decision threshold θ. Similarly, we define the empirical risk as $R_{emp}(\theta)$. The *generalization error bound* is an upper bound to the error $|R(\hat{\theta}) - R_{emp}(\hat{\theta})|$, where $\hat{\theta}$ is the minimizer of the empirical loss, i.e., $\hat{\theta} := \arg\min_\theta R_{emp}(\theta)$. In words, the bound provides a (statistical) guarantee on the performance $R(\hat{\theta})$, when using the learned $\hat{\theta}$ on unseen data, relative to the performance $R_{emp}(\hat{\theta})$ assessed on the training data. Our objective is to characterize this bound under censored feedback, and to evaluate how utilizing (pure or bounded) exploration can improve the bound.

Recall that the decision maker starts with a training data containing n_y IID samples from each label y, drawn from an underlying distribution with CDF $F^y(x)$. Let $n = n_0 + n_1$ denote the size of the initial training data. Then, the expected loss of a binary classifier with decision threshold θ is given by,

$$R(\theta) = \mathbb{E}_{XY}\mathcal{L}(f(X), Y) = p_1 F^1(\theta) + p_0(1 - F^0(\theta)) \ ,$$

while the empirical loss $R_{emp}(\theta)$ is given by,

$$R_{emp}(\theta) = \frac{n_1}{n}\frac{1}{n_1}\sum_{(x_i,y_i)} \mathbb{1}\{x_i \le \theta, y_i = 1\} + \frac{n_0}{n}\Big(1 - \frac{1}{n_0}\sum_{(x_i,y_i)} \mathbb{1}\{x_i \le \theta, y_i = 0\}\Big).$$

Similarly, if the decision maker can collect an additional k_y samples of agents with features above the threshold θ, the above empirical risk expression can be updated accordingly, by considering the $n_y + k_y$ samples available from each label y.

We detail the derivations of these expressions in Appendix I. Using these expressions of the expected and empirical risks, the following theorem provides an upper bound on the generalization error $|R(\hat{\theta}) - R_{emp}(\hat{\theta})|$ as a function of the CDF error bound, where $\hat{\theta}$ denotes the minimizer of the empirical loss, i.e., $\hat{\theta} := \arg\min_\theta R_{emp}(\theta)$.

Theorem 4. *Consider a threshold-based classifier $f_{\hat{\theta}}(x) : \mathcal{X} \to \{0, 1\}$, determined from a dataset containing n_y initial IID training samples from each label y, with $n = n_0 + n_1$, under a 0-1 loss function. Let p_y denote the proportion of agents from label y. Subsequently, due to the censored feedback, the algorithm collects k_y additional samples from each label y. Let F^y and F^y_m denote the CDFs and empirical CDFs, respectively, given m samples from label y agents. Then, with probability at least $1 - 2\delta$,*

$$\Big|R(\hat{\theta}) - R_{emp}(\hat{\theta})\Big| \le 3\Big|p_0 - \frac{n_0}{n}\Big| + \sum_{y\in\{0,1\}} \min\Big(p_y, \frac{n_y}{n}\Big) \sup_\theta \Big|F^y(\theta) - F^y_{n_y+k_y}(\theta)\Big| \ .$$

The proof is given in Appendix H. First, we note that tightening the CDF error bounds leads to tightening the generalization error guarantees. More specifically, using this theorem together with Theorems 1, 2, and 3, we can provide a generalization error guarantee for an algorithm in terms of the number of available data samples in its training data from each label and in different parts of the data domain, particularly when future data availability is non-IID due to censored feedback.

For instance, the DKW inequality can be alternatively expressed as follows: given n_y IID samples from a label y, with probability at least $1 - \delta$, the following inequality holds:

$$\sup_z \Big|F(z) - F^y_{n_y}(z)\Big| \le \sqrt{\frac{\log\frac{2}{\delta}}{2n_y}} \ .$$

Using this expression in Theorem 4, we conclude that (without censored feedback, or with pure exploration with $\epsilon = 1$) with probability at least $1 - 2\delta$,

$$\Big|R(\hat{\theta}) - R_{emp}(\hat{\theta})\Big| \le 3\Big|p_0 - \frac{n_0}{n}\Big| + \sum_{y\in\{0,1\}} \min\Big(p_y, \frac{n_y}{n}\Big)\sqrt{\frac{\log\frac{2}{\delta}}{2n_y}}.$$

We can similarly specialize Theorem 4 to tasks with censored feedback by linking it with Theorems 2 and 3. Given the complexity of the CDF error bounds under censored feedback, while we cannot derive a closed-form expression for the bound as done for the DKW inequality, we can compute the bounds numerically, as shown in the next section.

5 Numerical Experiments

5.1 CDF error bounds

We first illustrate our derived bounds (with $\delta = 0.015$) on the empirical CDF. We start with 50 random samples from a Gaussian distribution N(7,1). Next, 200 new samples are drawn from the same distribution, with all samples with features $x \geq \theta = 7$ accepted, and samples with features $LB = 6 \leq x \leq \theta$ accepted with a probability $\epsilon \in \{0, 0.5, 1\}$; higher values of ϵ the represent less censored feedback ($\epsilon = 1$ means no censored feedback).

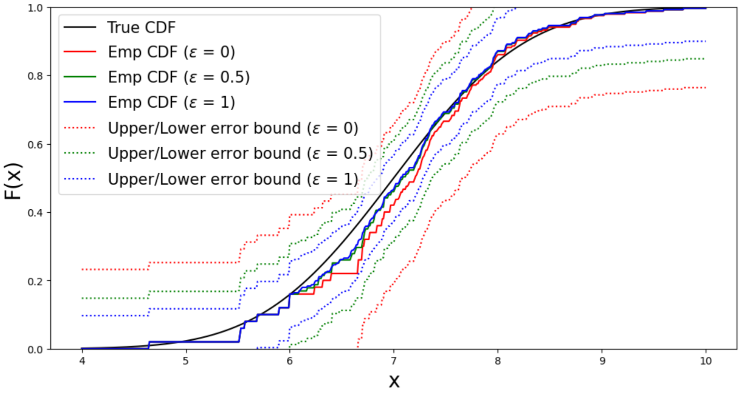

Figure 4: CDF error bounds when different levels of exploration (ϵ) are used to alleviate censored feedback. As ϵ increases: (a) the empirical CDF estimates become more accurate, and (b) our CDF error bounds improve (i.e., more tightly enclose the true CDF).

From Figure 4, we first note that our bounds (the dotted lines) effectively enclose the true distribution. We also note the distinction between empirical CDFs in the disclosed region ($x \geq 7$) and the censored region ($x \leq 7$): as intuitively expected, empirical CDFs (solid lines) in the disclosed region are "smoother" compared to those in the censored region. Furthermore, as ϵ (exploration) increases, we overcome censored feedback in the exploration region, resulting in more accurate empirical estimates. Additionally, as ϵ increases, our error bounds improve (i.e., more tightly enclose the true CDF).

5.2 Model generalization error bounds: real-world data and adaptively updated algorithm

We now illustrate the ability of our generalization error bounds (derived in Theorem 4) in providing guarantees on the error of the learned models from data affected by censored feedback, using experiments on a real-world dataset. In addition, while our bounds are derived for a fixed model $\hat{\theta}$, the model can be updated as new samples are collected. Therefore, in these experiments we also assess the performance of our bounds based on whether we adaptively update the decision threshold $\hat{\theta}$ with new samples.

We conduct these experiments on the real-world *Adult* census dataset (Dua & Graff, 2017). The objective is to predict whether an individual earns more than \$50k/year, based on a multi-dimensional feature set. We employ a logistic regression algorithm and 0-1 loss for the classification task, and compare the generalization error across different exploration probabilities ($\epsilon = \{0.5, 1\}$). We start with a 1000 sample training dataset. A total of 45000 new samples arrive throughout the experiment; in addition to accepting all samples with feature $x \geq \hat{\theta}$, the algorithm also accepts some samples that fall below $\hat{\theta}$. The decision threshold is updated periodically based on new data (after each 5000 batch of new samples arrives). The decision threshold is retrained using (most recent) training data. We report our experiment results for an average of 5 runs, where the randomness comes from the order of samples arrived and the exploration.

From Figure 5, we observe that as the decision threshold $\hat{\theta}$ is adaptively updated when more samples are collected, it has even better generalization performance compared to a non-adaptive decision threshold. This is expected as a refined decision threshold yields better performance on unseen data. Further, for the generalization error bounds (dotted lines in the right panel), we see that our bounds effectively contain the true generalization errors of the model (for both the fixed model and adaptively updated model cases). Notably, in the presence of censored feedback, we observe that the generalization error bound with adaptive updating is tighter than the non-adaptive one, pointing to a potential future research direction for further improving our bounds.

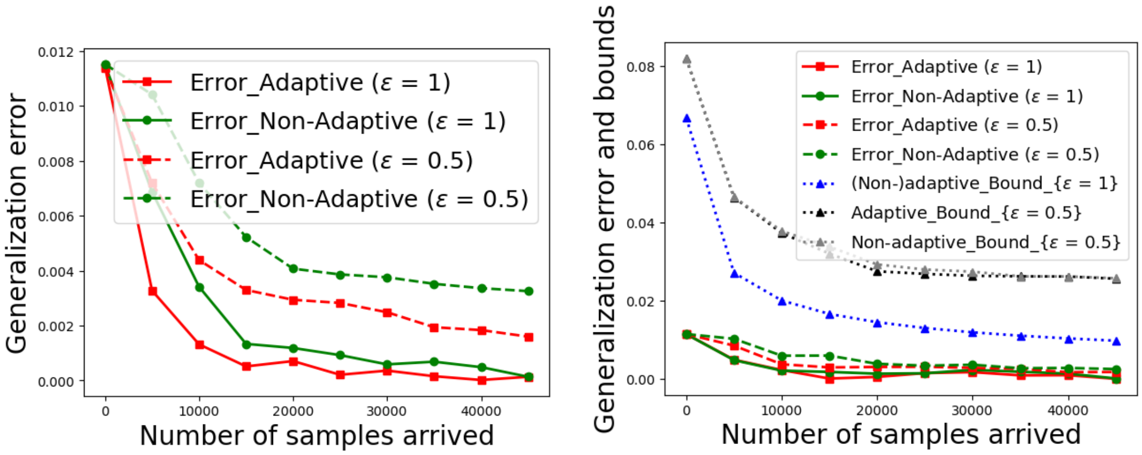

Figure 5: Generalization error with(out) an adaptively updated model ($\hat{\theta}$) and varying exploration (ϵ).

5.3 Comparison with existing generalization error bounds

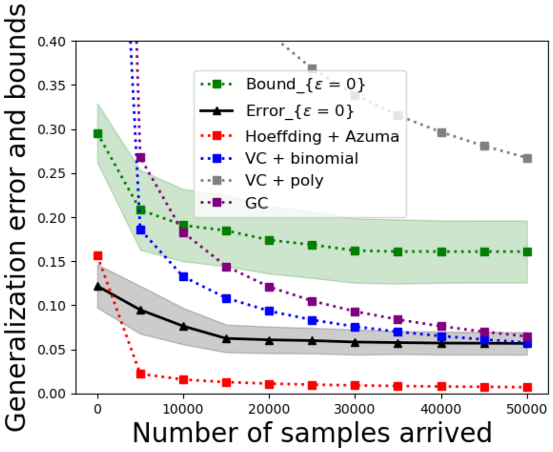

Figure 6: Existing bounds fail to capture generalization when there is censored feedback.

We now compare the performance of our bounds with a number of existing generalization error bounds, and show that by failing to account for censored feedback, prior works fail to correctly capture how well a model learned on data suffering from censored feedback generalizes to unseen data. We consider the following four benchmarks: The 'Hoeffding + Azuma' bounds represent those derived from Hoeffding and Azuma inequalities (Hoeffding, 1994; Azuma, 1967). The 'VC + binomial' bounds are VC generalization bounds (Vapnik & Chervonenkis, 2015; Abu-Mostafa et al., 2012, Thm 2.5) where the shatter coefficient is bounded through the binomial theorem. The 'VC + poly' bounds represent VC generalization bounds (Vapnik & Chervonenkis, 2015; Devroye et al., 2013, Thm 13.11) applicable to any linear classifier whose empirical error is minimal, where the shatter coefficient is bounded by a polynomial function. Lastly, the 'GC' bounds (Glivenko, 1933; Cantelli, 1933) are derived based on the Glivenko-Cantelli Theorem for a threshold classifier and 0-1 loss.

We conduct this experiment on synthetic data. We start with 50 initial training samples for each label $y \in \{0,1\}$ randomly drawn from Gaussian distributions N(9,1) and N(10,1), respectively. The decision threshold $\hat{\theta}$ is selected to be the one minimizing the misclassification error on the training data. Then, a total of 50000 new samples arrive throughout the experiment. They will be accepted if the feature $x \geq \hat{\theta}$, otherwise, they are rejected. We run the experiments 5 times and report the average results with corresponding error bars. From Figure 6(a), we can clearly see that the 'Hoeffding-Azuma' (red), 'VC+binomial' (blue), and 'GC' (purple) bounds are inadequate for accurately estimating the true generalization error guarantees of the model. For the 'VC+poly' (gray) bound, for the given number of new samples, it provides a very loose bound, even compared with our bounds. However, as the number of arrived samples increases, it will exhibit similar behaviors to the other three benchmarks, in that it will go lower than the true generalization error (black line/shades).

6 Conclusion and Future Work

We studied generalization error bounds for classification models learned from non-IID data collected under censored feedback. We presented two generalizations of the Dvoretzky-Kiefer-Wolfowitz (DKW) inequality, which characterizes the gap between empirical and theoretical CDFs given *IID* data, to problems with *non-IID* data due to censored feedback without exploration (Theorem 2) and with exploration (Theorem 3), and connected these bounds to generalization error guarantees of the learned model (Theorem 4). Our findings establish the extent to which a decision maker should be concerned about censored feedback's impact on the learned model's performance guarantees, and show that a minimum level of exploration is needed to alleviate it.

For future work, we are interested in strengthening our bounds by allowing the model (θ) to be adaptively updated as new samples are collected; as noted in Section 5, this could help further strengthen our error bounds. Generalization error bounds under a combination of censored feedback and domain adaptation are also worth exploring, wherein the initial training data distribution differs from the target domain distribution. Finally, we have provided extensions of the DKW inequality, which strengthens the VC inequality when data is real-valued, under censored feedback; providing similar extensions of the VC inequality for *multi-dimensional data* could be an interesting direction of future work. We discuss some initial findings and potential challenges of this extension below.

Bounds for higher dimensional data. When assessing generalization error under censored feedback in higher dimensional data, one approach could be to first reduce the dimensionality, enabling direct application of our findings. For instance, we have performed a mapping of multi-dimensional features to a single-dimensional representation in our experiments on the real-world *Adult* census dataset. However, this reduction may lead to some loss of information, potentially impacting algorithm performance. An alternative would be to follow our approach of identifying IID subspaces in the higher-dimensional data space, apply a *multivariate* DKW inequality (e.g., (Naaman, 2021)) in these subspaces, and then identify the appropriate error coefficients to re-assemble the subdomain bounds and find a CDF error bound for the entire data domain. We provide an analysis for 2D spaces based on this approach in Appendix J. A main challenge when doing so is that while the decision boundary can be any arbitrary line (determining the two subspaces in which data can be viewed as IID), the standard joint CDF calculates the probability that $X \leq x$ and $Y \leq y$, where x and y are vertical and horizontal cutoff values. To circumvent this mismatch, we start with an *adjusted* CDF which measures data density and counts existing vs. newly collected samples in a "rotated" data space, and subsequently map the CDF error bound of the adjusted CDF to a CDF error bound for the standard CDF (as detailed in Appendix J). Alternative error bounds that build on the VC inequality for multi-dimensional data (instead of multi-dimensional DKW inequalities), remain as a potential direction for future work.

References

Jacob D Abernethy, Kareem Amin, and Ruihao Zhu. Threshold bandits, with and without censored feedback. *Advances In Neural Information Processing Systems*, 29, 2016.

Yaser S. Abu-Mostafa, Malik Magdon-Ismail, and Hsuan-Tien Lin. *Learning From Data*. AMLBook, 2012.

Kazuoki Azuma. Weighted sums of certain dependent random variables. *Tohoku Mathematical Journal, Second Series*, 19(3):357–367, 1967.

Maria-Florina Balcan, Andrei Broder, and Tong Zhang. Margin based active learning. In *Learning Theory: 20th Annual Conference on Learning Theory, COLT 2007, San Diego, CA, USA; June 13-15, 2007. Proceedings 20*, pp. 35–50. Springer, 2007.

Peter L Bartlett and Shahar Mendelson. Rademacher and gaussian complexities: Risk bounds and structural results. *Journal of Machine Learning Research*, 3(Nov):463–482, 2002.

Yahav Bechavod, Katrina Ligett, Aaron Roth, Bo Waggoner, and Steven Z Wu. Equal opportunity in online classification with partial feedback. *Advances in Neural Information Processing Systems*, 32, 2019.

Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. A theory of learning from different domains. *Machine learning*, 79:151–175, 2010.

D Bitouzé, B Laurent, and Pascal Massart. A dvoretzky–kiefer–wolfowitz type inequality for the kaplan–meier estimator. In *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, volume 35, pp. 735–763. Elsevier, 1999.

Olivier Bousquet and André Elisseeff. Stability and generalization. *The Journal of Machine Learning Research*, 2:499–526, 2002.

Francesco Paolo Cantelli. Sulla determinazione empirica delle leggi di probabilita. *Giorn. Ist. Ital. Attuari*, 4(421-424), 1933.

Bowen Cheng, Yunchao Wei, Jiahui Yu, Shiyu Chang, Jinjun Xiong, Wen-Mei Hwu, Thomas S Huang, and Humphrey Shi. A simple non-iid sampling approach for efficient training and better generalization. *arXiv preprint arXiv:1811.09347*, 2018.

Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. Algorithmic decision making and the cost of fairness. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, pp. 797–806, 2017.

Corinna Cortes, Giulia DeSalvo, Claudio Gentile, Mehryar Mohri, and Ningshan Zhang. Region-based active learning. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 2801–2809. PMLR, 2019.

Corinna Cortes, Giulia DeSalvo, Claudio Gentile, Mehryar Mohri, and Ningshan Zhang. Adaptive region-based active learning. In *International Conference on Machine Learning*, pp. 2144–2153. PMLR, 2020.

Yash Deshpande, Lester Mackey, Vasilis Syrgkanis, and Matt Taddy. Accurate inference for adaptive linear models. In *International Conference on Machine Learning*, pp. 1194–1203. PMLR, 2018.

Luc Devroye, László Györfi, and Gábor Lugosi. *A probabilistic theory of pattern recognition*, volume 31. Springer Science & Business Media, 2013.

Dheeru Dua and Casey Graff. UCI machine learning repository, 2017. URL `http://archive.ics.uci.edu/ml`.

Aryeh Dvoretzky, Jack Kiefer, and Jacob Wolfowitz. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, pp. 642–669, 1956.

Danielle Ensign, Sorelle A Friedler, Scott Neville, Carlos Scheidegger, and Suresh Venkatasubramanian. Runaway feedback loops in predictive policing. In *Conference on fairness, accountability and transparency*, pp. 160–171. PMLR, 2018.

Valery Glivenko. Sulla determinazione empirica delle leggi di probabilita. *Gion. Ist. Ital. Attauri.*, 4:92–99, 1933.

Yair Goldberg. Hoeffding-type and bernstein-type inequalities for right censored data. *arXiv preprint arXiv:1903.01991*, 2019.

Yair Goldberg and Michael R Kosorok. Support vector regression for right censored data. 2017.

Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *The collected works of Wassily Hoeffding*, pp. 409–426, 1994.

Abbas Kazerouni, Qi Zhao, Jing Xie, Sandeep Tata, and Marc Najork. Active learning for skewed data sets. *arXiv preprint arXiv:2005.11442*, 2020.

Niki Kilbertus, Manuel Gomez Rodriguez, Bernhard Schölkopf, Krikamol Muandet, and Isabel Valera. Fair decisions despite imperfect predictions. In *International Conference on Artificial Intelligence and Statistics*, pp. 277–287. PMLR, 2020.

Aryeh Kontorovich and Roi Weiss. Uniform chernoff and dvoretzky-kiefer-wolfowitz-type inequalities for markov chains and related processes. *Journal of Applied Probability*, 51(4):1100–1113, 2014.

Vitaly Kuznetsov and Mehryar Mohri. Generalization bounds for non-stationary mixing processes. *Machine Learning*, 106(1):93–117, 2017.

Cheolhei Lee, Kaiwen Wang, Jianguo Wu, Wenjun Cai, and Xiaowei Yue. Partitioned active learning for heterogeneous systems. *Journal of Computing and Information Science in Engineering*, 23(4):041009, 2023.

Pascal Massart. The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *The annals of Probability*, pp. 1269–1283, 1990.

Dharmendra S Modha and Elias Masry. Minimum complexity regression estimation with weakly dependent observations. *IEEE Transactions on Information Theory*, 42(6):2133–2145, 1996.

Mehryar Mohri and Afshin Rostamizadeh. Stability bounds for non-iid processes. *Advances in Neural Information Processing Systems*, 20, 2007.

Mehryar Mohri and Afshin Rostamizadeh. Rademacher complexity bounds for non-iid processes. *Advances in Neural Information Processing Systems*, 21, 2008.

Michael Naaman. On the tight constant in the multivariate dvoretzky–kiefer–wolfowitz inequality. *Statistics & Probability Letters*, 173:109088, 2021.

Xinkun Nie, Xiaoying Tian, Jonathan Taylor, and James Zou. Why adaptively collected data have negative bias and how to correct for it. In *International Conference on Artificial Intelligence and Statistics*, pp. 1261–1269. PMLR, 2018.

David Pollard. *Convergence of stochastic processes*. Springer Science & Business Media, 2012.

Reilly Raab and Yang Liu. Unintended selection: Persistent qualification rate disparities and interventions. *Advances in Neural Information Processing Systems*, 34:26053–26065, 2021.

Steve Smale and Ding-Xuan Zhou. Online learning with markov sampling. *Analysis and Applications*, 7(01): 87–113, 2009.

Ingo Steinwart and Andreas Christmann. Fast learning from non-iid observations. *Advances in neural information processing systems*, 22, 2009.

Ingo Steinwart, Don Hush, and Clint Scovel. Learning from dependent observations. *Journal of Multivariate Analysis*, 100(1):175–194, 2009.

Xueyang Tang, Song Guo, and Jingcai Guo. Personalized federated learning with clustered generalization. 2021.

Vladimir N Vapnik and A Ya Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Measures of complexity: festschrift for alexey chervonenkis*, pp. 11–30, 2015.

Jing Wang, Laurel Hopkins, Tyler Hallman, W Douglas Robinson, and Rebecca Hutchinson. Cross-validation for geospatial data: Estimating generalization performance in geostatistical problems. *Transactions on Machine Learning Research*, 2023.

Dennis Wei. Decision-making under selective labels: Optimal finite-domain policies and beyond. In *International Conference on Machine Learning*, pp. 11035–11046. PMLR, 2021.

Yifan Yang, Yang Liu, and Parinaz Naghizadeh. Adaptive data debiasing through bounded exploration. *Advances in Neural Information Processing Systems*, 35:1516–1528, 2022.

Bin Yu. Rates of convergence for empirical processes of stationary mixing sequences. *The Annals of Probability*, pp. 94–116, 1994.

Zhilin Zhao, Longbing Cao, and Chang-Dong Wang. Gray learning from non-iid data with out-of-distribution samples. *arXiv preprint arXiv:2206.09375*, 2022.

Guanhua Zheng, Jitao Sang, Houqiang Li, Jian Yu, and Changsheng Xu. A generalization theory based on independent and task-identically distributed assumption. *arXiv preprint arXiv:1911.12603*, 2019.

Bin Zou, Luoqing Li, and Zongben Xu. The generalization performance of erm algorithm with strongly mixing observations. *Machine learning*, 75(3):275–295, 2009.

A Additional and detailed related work

Censored feedback (Adaptive sampling). The problem of censored feedback has been extensively studied in the literature. Prior works, such as (Abernethy et al., 2016; Deshpande et al., 2018; Ensign et al., 2018; Wei, 2021), have explored various aspects of this problem. For instance, Ensign et al. (2018) view the problem as a semi-bandit problem and apply existing regret bounds to obtain their theoretical result. Abernethy et al. (2016) propose a Kaplan-Meier-based Upper Confidence Bound algorithm for censored feedback in the context of a threshold bandit problem. Deshpande et al. (2018) investigate inference in high-dimensional linear models and analyze the sample adaptivity from a bias and variance trade-off perspective in both time series analysis and batched data collection. Wei (2021) studies the learning policies with censored feedback, finding the optimal decision policy that maximizes the utility function. This work considers both adaptive sampling and cost exploration and formulates the problem as a partially observed Markov decision process, which is solved through dynamic programming. Our work is similar to these prior works in that we also consider censored feedback, but we mainly focus on deriving a generalization error bound in the presence of censored feedback, and we propose to use a bounded exploration technique to improve the bounds.

Bandit literature. Our work is related to the bandit problem (Abernethy et al., 2016; Ensign et al., 2018), which deals with the exploration and exploitation dilemma, where the decision maker can obtain additional information to improve the generalization performance through exploration, while at the same time incurring a large risk of exploration. In the bandit problem, the decision maker explores other arms in various ways, such as randomly using the ϵ-greedy algorithm, by some form of highest uncertainty as in UCB algorithm, or by importance sampling approaches as in EXP3, etc. The key difference in our approach is that we consider bounded exploration (Balcan et al., 2007; Lee et al., 2023; Wei, 2021; Yang et al., 2022), where the lower bound is set because the cost of wrong decisions increases as samples further away from the current decision threshold are admitted. Furthermore, in the bandit literature, regret analysis is conducted to analyze the model performance compared to the best actions. In contrast, we analyze the model performance from different angles, such as how to improve the generalization performance (upper bound on the difference between the model's performance on training data and unseen testing data) by utilizing the newly collected samples through exploration.

Generalization with non-IIDness. In addition to the β-mixing sequence non-IIDness discussed in the main paper, (Steinwart & Christmann, 2009; Steinwart et al., 2009) have studied a more general non-IID setting called α-mixing sequence, where stationarity is not necessary. Furthermore, (Modha & Masry, 1996; Zou et al., 2009) have explored the generalization performance based on exponentially strongly mixing observations, where the α-mixing coefficient is bounded by an exponential decay term. Apart from the mixing sequence, (Cheng et al., 2018) have imitated human learning behavior by discarding training samples with lower losses during training, and refreshing them when a new training epoch starts. They experimentally show that their method of using non-IID samples can lead to better performance than uniformly sampling using IID samples. (Smale & Zhou, 2009) have investigated non-IIDness in online learning algorithms through a Markov sampling method, where samples are drawn according to a sequence of probability measures. Meanwhile, (Zhao et al., 2022) explore non-IIDness through a mixing of in-distribution and out-of-distribution samples. They show that the performance gap is related to the distribution discrepancy between in-distribution and out-of-distribution (similar to (Ben-David et al., 2010), which refers to source and target discrepancy), empirical losses, and empirical estimation error.

DKW-typed inequalities with non-IIDness. Our work is also related to DKW-type inequalities with non-IID data. (Kontorovich & Weiss, 2014) derive deviation bounds using DKW-type inequalities for strongly mixing non-stationary Markov chains. Extending the DKW inequality with one-dimensional IID samples sharpened by (Massart, 1990), (Naaman, 2021) generalize it to the multi-dimensional case, incorporating an α-mixing condition. Furthermore, DKW-type inequalities have been applied to the Kaplan-Meier (KM) estimator in the analysis of right censored data in survival analysis, as shown by (Bitouzé et al., 1999). They utilize a weight factor, similar to our scaling factor, to control uniform deviation in the uncensored data model. Moreover, (Goldberg & Kosorok, 2017) introduce a censored SVM learning algorithm with a inverse probability censoring weighting (IPCW) loss function. Additionally, (Goldberg, 2019) develope Hoeffding-type and Bernstein-type inequalities based on DKW inequalities to bound the difference between the IPCW estimator and its expectation.

Vapnik-Chevronenkis (VC) Theory. In binary classification with the 0-1 loss function, the VC theory (Pollard, 2012, Thm 12.6) provides a bound on the maximum difference between the empirical estimates of whether an event happens and its expectation over IID data. It states that the difference could be bounded by $8S(\mathcal{F}, n)\exp(-n\eta^2/32)$, where $S(\mathcal{F}, n)$ is the shattering coefficient of $\mathcal{F}$ and n is the sample size. The shattering coefficient satisfies $S(\mathcal{F}, n) \leq (n+1)^d$ with finite VC dimension d so that the generalization error of linear classifiers with one dimensional IID samples could be bounded as $8(n+1)^2\exp(-n\eta^2/32)$. Alternatively, as the VC theory is a mighty generalization of Gilvenko-Cantelli (GC) bound (Pollard, 2012, Thm 12.4), when considering the deviation between the empirical and true CDFs, it takes the form of $8(n+1)\exp(-n\eta^2/32)$ by replacing $S(\mathcal{F}, n)$ with $n+1$. In our study, we employ the DKW bound, which is an extension of the GC bound by eliminating the logarithmic term $(n+1)$. The main distinction between our bounds and the VC/GC bounds lies in the assumption that the data input is non-IID over the entire data domain, while it may be treated as IID within each subdomain.

B Proof of Lemma 1

Proof. Let G and G_m be the theoretical and empirical CDFs of Z in the censored region, respectively.

For the full c.d.f. in $(-\infty, \theta)$, we have

$$F(x) = Pr(X \leq x), \qquad F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i \leq x}, \quad \forall i, x \in (-\infty, \theta)$$

For the censored c.d.f. in $(-\infty, \theta)$, we have

$$G(x) = Pr(Z \leq x), \qquad G_m(x) = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}_{Z_i \leq x}, \quad \forall i, x \in (-\infty, \theta)$$

Hence, for any fixed $x \in (-\infty, \theta)$, we have

$$\sum_{i=1}^{m} \mathbf{1}_{Z_i \leq x} = \sum_{i=1}^{n} \mathbf{1}_{X_i \leq x} = A$$

and

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i \leq x} = \frac{1}{n}A = \frac{m}{n}\cdot\frac{1}{m}A = \frac{m}{n}G_m(x)$$

Similarly,

$$G(x) = Pr(Z \leq x) = Pr(X \leq x | X \leq \theta) = \frac{Pr(X \leq x)}{Pr(X \leq \theta)} = \frac{F(x)}{\alpha}$$

Therefore, we have

$$\begin{aligned}
\sup_{x\in(-\infty,\theta)} |F - F_n| &= \sup_{x\in(-\infty,\theta)} \left|\alpha G - \frac{m}{n}G_m\right| \\
&= \sup_{x\in(-\infty,\theta)} \left|\alpha G - \alpha G_m + \alpha G_m - \frac{m}{n}G_m\right| \\
&\leq \sup_{x\in(-\infty,\theta)} \left|\alpha(G - G_m)\right| + \sup_{x\in(-\infty,\theta)} \left|(\alpha - \frac{m}{n})G_m\right| \\
&= \sup_{x\in(-\infty,\theta)} \left|\alpha(G - G_m)\right| + \left|\alpha - \frac{m}{n}\right|
\end{aligned}$$

where the last equality holds since $|\alpha - \frac{m}{n}|$ is constant once θ is fixed and by the monotonic property of the CDF curve. Notice that, by the definition of $G_m = \frac{1}{m}\sum_i \mathbf{1}_{Z_i \leq \theta} = 1$. We could instead choose to add and subtract $\frac{m}{n}G$, which would yield a similar result. Choosing the smaller of the two gives us the bound. ☐

C Proof of Lemma 2

Proof. Let K and K_{n-m+k} be the theoretical and empirical CDFs of Z in the disclosed region, respectively.

$$K(x) = \mathbb{P}(Z \leq x), \quad K_{n-m+k}(x) = \frac{1}{n-m+k} \sum_{i=1}^{n-m+k} \mathbb{1}(Z_i \leq x), \quad \forall x \in (\theta, \infty).$$

Since $Z = \{X | X \geq \theta\}$, then we have

$$K(x) = Pr(Z \leq x) = Pr(X \leq x | X \geq \theta) = \frac{Pr(\theta \leq X \leq x)}{Pr(X \geq \theta)} = \frac{F(x) - \alpha}{1 - \alpha}$$

In terms of the empirical CDF, the full empirical CDF on variable X in (θ, ∞) can be viewed as starting at the $F_n(\theta) = \frac{m}{n}$, then increasing to 1 as x increases with the total increment $\frac{n-m}{n}$. However, for the increment, the only difference between F_{n+k} and K_{n-m+k} is that the earlier one focuses on the variable X while the latter one focuses on the variable $Z = \{X | X \geq \theta\}$. Note that, the notation F_{n+k}, similar to F_n, means the empirical CDF is evaluated using $n + k$ samples. Hence, with additional k samples, we can use a scaling factor $\frac{n-m}{n}$ to write the CDF such that

$$F_{n+k}(x) = \frac{m}{n} + \frac{n-m}{n} \frac{1}{n-m+k} \sum_{i=1}^{n-m+k} \mathbf{1}_{\theta \leq X_i \leq x}, \quad \forall i, x \in (\theta, \infty)$$

Then, for any fixed $x \in (\theta, \infty)$, we have

$$F_{n+k}(x) - \frac{m}{n} = \frac{n-m}{n} \frac{1}{n-m+k} \sum_{i=1}^{n-m+k} \mathbf{1}_{\theta \leq X_i \leq x} = \frac{n-m}{n} K_{n-m+k}(x)$$

Therefore, we have

$$\begin{aligned}
\sup_{x \in (\theta,\infty)} |F - F_{n+k}| &= \sup_{x \in (\theta,\infty)} \left| (1-\alpha)K + \alpha - \frac{n-m}{n} K_{n-m+k} - \frac{m}{n} \right| \\
&\leq \sup_{x \in (\theta,\infty)} \left| (1-\alpha)K - \frac{n-m}{n} K_{n-m+k} \right| + \left| \alpha - \frac{m}{n} \right| \\
&= \sup_{x \in (\theta,\infty)} \left| (1-\alpha)K - (1-\alpha)K_{n-m+k} + (1-\alpha)K_{n-m+k} - \frac{n-m}{n} K_{n-m+k} \right| + \left| \alpha - \frac{m}{n} \right| \\
&\leq \sup_{x \in (\theta,\infty)} \left| (1-\alpha)(K - K_{n-m+k}) \right| + 2 \left| \alpha - \frac{m}{n} \right|
\end{aligned}$$

Similarly, we could instead choose to add and subtract $\frac{n-m}{n} K$ and choose the smaller of the two. □

D Proof of Theorem 2

Proof.

$$
\begin{aligned}
&P\Big(\sup_{x\in R}\Big|F(x)-F_{n+k}(x)\Big|\geq\eta\Big)\\
&=P\Big(\max\Big(\sup_{x\in(-\infty,\theta)}\Big|F(x)-F_n(x)\Big|,\sup_{x\in(\theta,\infty)}\Big|F(x)-F_{n+k}(x)\Big|\Big)\geq\eta\Big)\\
&=P\Big(\sup_{x\in(-\infty,\theta)}\Big|F(x)-F_n(x)\Big|\geq\eta\text{ or }\sup_{x\in(\theta,\infty)}\Big|F(x)-F_{n+k}(x)\Big|\geq\eta\Big)\\
&\leq P\Big(\sup_{x\in(-\infty,\theta)}\Big|F(x)-F_n(x)\Big|\geq\eta\Big)+P\Big(\sup_{x\in(\theta,\infty)}\Big|F(x)-F_{n+k}(x)\Big|\geq\eta\Big)\\
&\leq P\Big(\sup_{x\in(-\infty,\theta)}\Big|\min\Big(\alpha,\frac{m}{n}\Big)(G-G_m)\Big|+\Big|\alpha-\frac{m}{n}\Big|\geq\eta\Big)\\
&\qquad+P\Big(\sup_{x\in(\theta,\infty)}\Big|\min\Big(1-\alpha,\frac{n-m}{n}\Big)(K-K_{n-m+k})\Big|+2\Big|\alpha-\frac{m}{n}\Big|\geq\eta\Big)\\
&=P\Big(\sup_{x\in(-\infty,\theta)}\Big|G-G_m\Big|\geq\frac{\eta-\Big|\alpha-\frac{m}{n}\Big|}{\min\Big(\alpha,\frac{m}{n}\Big)}\Big)\\
&\qquad+P\Big(\sup_{x\in(\theta,\infty)}\Big|K-K_{n-m+k}\Big|\geq\frac{\eta-2\Big|\alpha-\frac{m}{n}\Big|}{\min\Big(1-\alpha,\frac{n-m}{n}\Big)}\Big)\\
&\leq 2\exp\Big(\frac{-2m(\eta-|\alpha-\frac{m}{n}|)^2}{\min\Big(\alpha,\frac{m}{n}\Big)^2}\Big)+2\exp\Big(\frac{-2(n-m+k)(\eta-2|\alpha-\frac{m}{n}|)^2}{\min\Big(1-\alpha,\frac{n-m}{n}\Big)^2}\Big)
\end{aligned}
$$

where the first inequality is obtained by the union bound, the second inequality is obtained according to Lemma 1 and 2, and the last inequality is obtained by plugging in the DKW inequality introduced in Theorem 1. ☐

E Proof of Theorem 3

Proof.

$$
\begin{aligned}
&P\Big(\sup_{x\in R}\Big|F(x)-F_{n+k_1+k_2}(x)\Big|\geq\eta\Big)\\
&=P\Big(\sup_{x\in(-\infty,\mathrm{LB})}\Big|F(x)-F_n(x)\Big|\geq\eta\text{ or }\sup_{x\in(\mathrm{LB},\theta)}\Big|F(x)-F_{n+k_1}(x)\Big|\geq\eta\\
&\qquad\qquad\text{ or }\sup_{x\in(\theta,\infty)}\Big|F(x)-F_{n+k_1+k_2}(x)\Big|\geq\eta\Big)\\
&\leq P\Big(\sup_{x\in(-\infty,\mathrm{LB})}\Big|\min\Big(\beta,\frac{l}{n}\Big)(G-G_l)\Big|+\Big|\beta-\frac{l}{n}\Big|\geq\eta\Big)
\end{aligned}
$$

$$
\begin{aligned}
&+ P\Bigg(\sup_{x\in(\mathrm{LB},\theta)} \Big| \min\big(\alpha-\beta, \tfrac{n-l}{n} \tfrac{m-l+k_1}{n-l+k_1+\epsilon k_2}\big)(E - E_{m-l+k_1})\Big| \\
&\qquad\qquad + \Big|\alpha-\beta-\tfrac{n-l}{n}\tfrac{m-l+k_1}{n-l+k_1+\epsilon k_2}\Big| + \Big|\beta - \frac{l}{n}\Big| \geq \eta \Bigg) \\
&+ P\Bigg(\sup_{x\in(\theta,\infty)} \Big| \min\Big(1-\alpha, \tfrac{n-l}{n} \tfrac{n-m+\epsilon k_2}{n-l+k_1+\epsilon k_2}\Big)(K - K_{n-m+k_2})\Big| \\
&\qquad\qquad + 2\Big|\alpha - \tfrac{l}{n} - \tfrac{n-l}{n}\tfrac{m-l+k_1}{n-l+k_1+\epsilon k_2}\Big| \geq \eta \Bigg) \\
= {} & P\Bigg(\sup_{x\in(-\infty,\mathrm{LB})} \Big| G - G_l \Big| \geq \frac{\eta - \Big|\beta - \frac{l}{n}\Big|}{\min\Big(\beta, \frac{l}{n}\Big)} \Bigg) \\
&+ P\Bigg(\sup_{x\in(\mathrm{LB},\theta)} \Big| E - E_{m-l+k_1} \Big| \geq \frac{\eta - \Big|\beta - \frac{l}{n}\Big| - \Big|\alpha-\beta-\frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\Big|}{\min\Big(\alpha-\beta, \frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\Big)} \Bigg) \\
&+ P\Bigg(\sup_{x\in(\theta,\infty)} \Big| K - K_{n-m+k_2} \Big| \geq \frac{\eta - 2\Big|\alpha - \frac{l}{n} - \frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\Big|}{\min\Big(1-\alpha, \frac{n-l}{n}\frac{n-m+\epsilon k_2}{n-l+k_1+\epsilon k_2}\Big)} \Bigg) \\
\leq {} & 2\exp\Big(\frac{-2l(\eta - |\beta - \frac{l}{n}|)^2}{\min\Big(\beta, \frac{l}{n}\Big)^2}\Big) \\
&+ 2\exp\Big(\frac{-2(m-l+k_1)\Big(\eta - |\beta - \frac{l}{n}| - \Big|\alpha-\beta-\frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\Big|\Big)^2}{\min\Big(\alpha-\beta, \frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\Big)^2}\Big) \\
&+ 2\exp\Big(\frac{-2(n-m+k_2)(\eta - 2|\alpha - \frac{l}{n} - \frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}|)^2}{\min\Big(1-\alpha, \frac{n-l}{n}\frac{n-m+\epsilon k_2}{n-l+k_1+\epsilon k_2}\Big)^2}\Big)
\end{aligned}
$$

The proof is similar to the proof of Theorem 2, with the introduction of exploration region and re-estimated empirical estimates α. ☐

F Proof of Proposition 1

Proof. Before we dive into the proof of the proposition, we first show that the impact of the estimation error under different θ is minimal compared to the change of θ.

Let's consider the censored region (first term) in the above inequality, and w.l.o.g., we assume that $m/n < \alpha$ and $\alpha - m/n = u$. Then, the first term could be written as

$$2\exp\left(\frac{-2m(\eta-u)^2}{(\frac{m}{n})^2}\right) = 2\exp\left(\frac{-2n^2(\eta-u)^2}{m}\right)$$

As θ decreases, m decreases. It is worth noting that $E(u) = 0$ according to the law of large numbers. In the following analysis, we measure the change as factor b, b' for m, u, respectively, and denote Δ_m, Δ_u as the change in the output. We also assume $0 \leq b, b' \leq 1$, and use it as a proxy to show the decreasing in θ. In other words, we are analyzing the impact of change in u compared to the change in m as θ decreases, our objective is to show the following

$$\frac{\Delta_m}{\Delta_u} = \frac{2\exp\left(\frac{-2n^2(\eta-u)^2}{m}\right) - 2\exp\left(\frac{-2n^2(\eta-u)^2}{bm}\right)}{2\exp\left(\frac{-2n^2(\eta-u)^2}{m}\right) - 2\exp\left(\frac{-2n^2(\eta-b'u)^2}{m}\right)} > 1$$

which is equivalent to show

$$\frac{\exp\left(\frac{-2n^2(\eta-u)^2}{bm}\right)}{\exp\left(\frac{-2n^2(\eta-b'u)^2}{m}\right)} < 1$$

Since we have,

$$\begin{aligned}\frac{\exp\left(\frac{-2n^2(\eta-u)^2}{bm}\right)}{\exp\left(\frac{-2n^2(\eta-b'u)^2}{m}\right)} &= \exp(\frac{-2n^2(\eta-u)^2}{bm} - \frac{-2n^2(\eta-b'u)^2}{m}) \\ &= \exp(\frac{-\frac{2}{b}n^2(\eta-u)^2 + 2n^2(\eta-b'u)^2}{m}) \\ &= (\exp(-\frac{2n^2}{m}))^{(\frac{1}{b}(\eta-u)^2-(\eta-b'u)^2)}\end{aligned}$$

It is clear that the exponent is greater than 0 since the change magnitude of m could be much greater than that of u as θ decreases, meaning that $0 < b << b' \approx 1$. Therefore, the impact of the estimation error under different θ is minimal compared to the change of θ. As a result, the first term decreases as θ decreases, primarily due to the decrease in the denominator.

Regarding the disclosed region (second term), the sample size increases due to a reduction in the number of samples (m) in the censored region and the additional samples (k) collected from the disclosed region. Similar to the discussion in the proof of Proposition 1, let $k = c(n-m)$, where c is a scalar representing how many additional samples we could collect proportionally above the decision threshold. If $c \geq \frac{(n-m)(\eta-u)^2}{m(\eta-2u)^2} - 1$, as more additional samples are collected, it will contribute to making the second term vanish and become dominated by the first term. Consequently, as θ decreases, the bounds also decrease, leading to $B(\theta)$ decreases as θ decreases.

□

G Proof of Proposition 2

Proof. As ϵ decreases, the censored region (first term) remains unchanged since the change in ϵ only affects the value of k_1. For the disclosed region (third term), similar to Proposition 1, the impact of the estimation error $|\alpha - \frac{l}{n} - \frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}|$ under different values of ϵ is minimal compared to the change in k_1. Therefore, the whole expression is dominated by the change in the exploration region (second term).

For the exploration region, it is worth noting that when $\epsilon = 1$, we collect k_1 samples in the exploration region. Therefore, if $\epsilon < 1$, the number of k_1 will decrease leading to an increase in the exponential term. Thus, as ϵ becomes smaller, the bound in the exploration region increases, resulting in an overall increase in the bounds. This implies that $B^e(LB, \theta, \epsilon)$ is decreasing in ϵ.

□

H Proof of Theorem 4

Proof. Let $F^1_{n_1+k_1}(\theta) := \frac{1}{n_1}\sum_{(x_i,y_i)} \mathbb{1}\{x_i \leq \theta, y_i = 1\}$ and $F^0_{n_0+k_0}(\theta) := \frac{1}{n_0}\sum_{(x_i,y_i)} \mathbb{1}\{x_i \leq \theta, y_i = 0\}$ be the empirical estimate of the area of the censored region for the corresponding label.

$$\begin{aligned}
\Big|R(\hat{\theta}) - R_{emp}(\hat{\theta})\Big| &\leq \sup_{\theta}\Big|R(\theta) - R_{emp}(\theta)\Big| \\
&= \sup_{\theta}\Big|p_1 F^1(\theta) + p_0\Big(1 - F^0(\theta)\Big) - \frac{n_1}{n}F^1_{n_1+k_1}(\theta) - \frac{n_0}{n}\Big(1 - F^0_{n_0+k_0}(\theta)\Big)\Big| \\
&\leq \sup_{\theta}\Big|p_1 F^1(\theta) - \frac{n_1}{n}F^1_{n_1+k_1}(\theta)\Big| + \sup_{\theta}\Big|p_0 F^0(\theta) - \frac{n_0}{n}F^0_{n_0+k_0}(\theta)\Big| + |p_0 - \frac{n_0}{n}| \\
&\leq \sum_{y\in\{0,1\}} \min\Big(p_y, \frac{n_y}{n}\Big)\sup_{\theta}\Big|F^y(\theta) - F^y_{n_y+k_y}(\theta)\Big| + 3|p_0 - \frac{n_0}{n}|
\end{aligned}$$

The last inequality follows the same technique as the proof of Lemma B.

□

I Expected and empirical loss derivation

Following the setup described in Section 2, let F^y and F^y_n denote the CDFs and empirical CDFs, respectively, with n realizations of the distribution of label $y \in \{0,1\}$. We assume that among these n samples, n_1 of them belong to qualified agents ($y = 1$), while the remaining ones are from unqualified agents ($y = 0$). We suppose the model is equipped with a 0-1 loss function and a threshold-based binary classifier $f_\theta(x) : \mathcal{X} \to \{0,1\}$. We then can derive the expected and empirical loss as follows:

$$\begin{aligned}
R(\theta) &= E_{XY} l(f(X), Y) \\
&= Pr(Y = 1)\cdot E_{X|Y=1} l(f(X), Y = 1) + Pr(Y = 0)\cdot E_{X|Y=0} l(f(X), Y = 0) \\
&= p_1 E\Big[\mathbb{1}\{f(X) \neq Y, Y = 1\}\Big] + p_0 E\Big[\mathbb{1}\{f(X) \neq Y, Y = 0\}\Big] \\
&= p_1 Pr\Big(\mathbb{1}\{f(X) \neq Y, Y = 1\}\Big) + p_0 Pr\Big(\mathbb{1}\{f(X) \neq Y, Y = 0\}\Big) \\
&= p_1 Pr\Big(f(X) \leq \theta, Y = 1\Big) + p_0 Pr\Big(f(X) \geq \theta, Y = 0\Big) \\
&= p_1 F^1(\theta) + p_0(1 - F^0(\theta))
\end{aligned}$$

$$\begin{aligned}
R_{emp}(\theta) &= \frac{1}{n}\sum_{(x_i,y_i)} l(f(x_i), y_i) = \frac{1}{n}\sum_{(x_i,y_i)} \mathbb{1}\{f(x_i) \neq y_i\} \\
&= \frac{1}{n}\Big[\sum_{y_i=1} \mathbb{1}\{f(x_i) \neq 1\} + \sum_{y_i=0} \mathbb{1}\{f(x_i) \neq 0\}\Big] \\
&= \frac{n_1}{n}\frac{1}{n_1}\sum_{(x_i,y_i)} \mathbb{1}\{x_i \leq \theta, y_i = 1\} + \frac{n_0}{n}\Big(1 - \frac{1}{n_0}\sum_{(x_i,y_i)} \mathbb{1}\{x_i \leq \theta, y_i = 0\}\Big)
\end{aligned}$$

J Analysis with higher dimensional samples

In this section, we expand our analysis from real-valued data to a higher dimensional space using the multivariate DKW inequality introduced in (Naaman, 2021). We first state the inequality and subsequently show how we will split the data domain to accommodate censored feedback. Although this analysis applies to samples with $k > 1$ dimensions, for the sake of notation simplicity, we focus on samples with two dimensions.

Theorem 5 (Multivariate DKW inequality (Naaman, 2021))**.** *Let $Z_1, \ldots, Z_n$ be k dimensional IID random variables with the empirical distribution function be $F_n(z) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}(Z_i \leq z)$. Let $F(z)$ be the expectation of $F_n(z)$. Then, for every n and η> 0,*

$$\mathbb{P}\Big(\sup_{z\in\mathbb{R}^k}\Big|F(z) - F_n(z)\Big| \geq \eta\Big) \leq 2k\exp\left(-2n\eta^2\right) .$$

In the 2D space, the threshold classifier is a linear classifier, with the decision boundary (line) represented as ($\mathbf{W}^T\mathbf{X} - b = 0$). Here, the bold terms denote vector forms, where $\mathbf{W}^T = [w_1, w_2]$ and $\mathbf{X}^T = [X_1, X_2]$. This line splits the full data domain ($\mathbb{R}^2$) into two subspaces: the subspace where $\mathbf{W}^T\mathbf{X} - b < 0$ is the censored region, while the remaining half space is the disclosed region.

We first establish the relation between the deviation of the censored partial empirical CDF from its expectation, and the deviation of the full empirical CDF from its expectation. The proof mirrors that of Lemma 1.

Lemma 4 (Multivariate Censored Region)**.** *Let* $\boldsymbol{Z} = \{\boldsymbol{X} | \mathbf{W}^T\mathbf{X} - b < 0\}$ *denote the* m *out of* $n + k$ *samples that are in the censored region (C). Let* $\alpha = F(\mathbf{W}^T\mathbf{X} - b < 0)$*. Let* G *and* G_m *be the theoretical and empirical CDFs of* $\boldsymbol{Z}$*, respectively. Then,*

$$\sup_{\mathbf{x}\in C} |F(\mathbf{x}) - F_{n+k}(\mathbf{x})| \le \sup_{\mathbf{x}\in C} \underbrace{\left|\min\left(\alpha, \frac{m}{n}\right)(G(\mathbf{x}) - G_m(\mathbf{x}))\right|}_{\textit{scaled censored subdomain error}} + \underbrace{\left|\alpha - \frac{m}{n}\right|}_{\textit{scaling error}} .$$

In the disclosed region, we could derive the expression following a similar approach. However, in 2D space, for any given $\mathbf{x}$ in the disclosed region, the number of samples in the disclosed region enclosed by $F(\mathbf{x})$ should be no less than $n+k-m$ because we only need to subtract the number of samples from the area of $\mathbf{W}^T\mathbf{X}-b < 0$, $X_1 \le x_1$ and $X_2 \le x_2$ out of m. Since we only know there are a total of m samples located in the censored region, we introduce an adjusted CDF $\tilde{F}$, compared to the standard CDF $F(\mathbf{x}) = Pr(X_1 \le x_1, X_2 \le x_2)$.

Definition 1 (Adjusted CDF)**.** *Let the decision line* $\mathbf{W}^T\mathbf{X} - b = 0$ *split the entire data domain into two subspaces. Let the adjusted CDF be*

$$\tilde{F}(x_1, b') = \int_{-\infty}^{\infty} \int_{-\infty}^{-\frac{w_1}{w_2}x_1 + \frac{w_1}{w_2}b'} f_{x_1,x_2} d_{x_1} d_{x_2}$$

where w_1, w_2 *are the parameters defined by the line, and* $b' \in (-\infty, \infty)$*.*

Remark. The standard CDF F calculates the area of $(X_1 \le x_1, X_2 \le x_2)$, where x_1 and x_2 are horizontal and vertical thresholds, respectively. In contrast, the adjusted CDF measures the area below the line $\mathbf{W}^T\mathbf{X} = b'$. When $b' = -\infty$(resp. ∞), it is equivalent to the case of ($x_1 = x_2 = -\infty$(resp. ∞)). Similarly, when $b' = b$, $\tilde{F}(x_1, b' = b) = F(\mathbf{W}^T\mathbf{X} - b < 0) = \alpha$. The adjusted CDF and standard CDF are both non-decreasing functions ranging from 0 to 1, and they only differ in the ways of calculating the enclosed area. By using adjusted CDFs as a proxy, we can derive such relations in the disclosed region.

Lemma 5 (Multivariate Disclosed Region)**.** *Let* $\boldsymbol{Z} = \{\boldsymbol{X} | \mathbf{W}^T\mathbf{X} - b > 0\}$ *denote the* $n - m + k$ *out of the* $n + k$ *samples in the disclosed region (D). Let* $\alpha = F(\mathbf{W}^T\mathbf{X} - b < 0)$*. Let* $\tilde{K}$ *and* $\tilde{K}_{n-m+k}$ *be the adjusted theoretical and empirical CDFs of* $\boldsymbol{Z}$*, respectively. Then,*

$$\sup_{\mathbf{x}\in D} |F(\mathbf{x}) - F_{n+k}(\mathbf{x})| \le \sup_{\mathbf{x}\in D} \underbrace{\left|\min(1-\alpha, 1 - \frac{m}{n})(\tilde{K}(\mathbf{x}) - \tilde{K}_{n-m+k}(\mathbf{x}))\right|}_{\textit{scaled disclosed subdomain error}} + \underbrace{2\left|\alpha - \frac{m}{n}\right|}_{\textit{shifting and scaling errors}}$$

Proof. Given any $\mathbf{x} = (x_1, x_2)$, we can always find a line passing through the point (x_1, x_2) with the same $\mathbf{W}$ as the decision line. Moreover, the area below the line is greater than that measured by the standard CDF. Since our focus is on $\sup_{\mathbf{x}} |F(\mathbf{x}) - F_{n+k}(\mathbf{x})|$, the maximum differences between empirical and theoretical CDFs can be upper bounded in a larger area (e.g., the area below the line). Therefore, we can obtain the following inequality:

$$\sup_{\mathbf{x}\in D} |F(\mathbf{x}) - F_{n+k}(\mathbf{x})| \le \sup_{\mathbf{x}\in D} |\tilde{F}(\mathbf{x}) - \tilde{F}_{n+k}(\mathbf{x})|$$

The equality holds when the maximum difference occurs in the area of $(X_1 \le x_1, X_2 \le x_2)$. The remainder of the proof follows that of Lemma 2. ☐

Now, we present our theorem, extending the multivariate DKW inequality to problems with censored feedback. Analogous to Theorem 2, we derive the following based on Lemma 4 and 5:

Theorem 6. *Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ be initial/historical IID data samples with cumulative distribution function $F(\boldsymbol{X})$. Let $\mathbf{W}^T\mathbf{X} - b = 0$ partition the data domain into two regions, such that $\alpha = F(\mathbf{W}^T\mathbf{X} - b < 0)$, and m of the initial n samples are located in the censored region. Assume we collect k additional samples from the disclosed region, and let F_{n+k} denote the empirical CDF estimated from these $n + k$ (non-IID) data. Then, for every m, n, k and $\eta > 0$,*

$$\mathbb{P}\left[\sup_{\boldsymbol{x}\in\mathbb{R}^2}\left|F(\boldsymbol{x}) - F_{n+k}(\boldsymbol{x})\right| \geq \eta\right] \leq \underbrace{4\exp\left(\frac{-2m(\eta-|\alpha-\frac{m}{n}|)^2}{\min\left(\alpha, \frac{m}{n}\right)^2}\right)}_{\textit{censored region error (constant)}} + \underbrace{4\exp\left(\frac{-2(n-m+k)(\eta-2|\alpha-\frac{m}{n}|)^2}{\min\left(1-\alpha, \frac{n-m}{n}\right)^2}\right)}_{\textit{disclosed region error (decreasing with additional data)}}$$

Similarly, incorporating the exploration, the LB can now be defined as a new line sharing the same $\mathbf{W}$ as the decision line but with a different intercept (e.g., $b_{\mathrm{LB}} < b$). Consequently, we have the following:

Theorem 7. *Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ be initial/historical IID data samples with cumulative distribution function $F(\boldsymbol{X})$. Let $\mathbf{W}^T\mathbf{X} - b_{LB} = 0$ and $\mathbf{W}^T\mathbf{X} - b = 0$ partition the data domain into three regions, such that $\beta = F(\mathbf{W}^T\mathbf{X} - b_{LB} < 0)$ and $\alpha = F(\mathbf{W}^T\mathbf{X} - b < 0)$, with l and m of the initial n samples located in the censored region, and the combined censored and exploration region, respectively. Assume we collect an additional k_1 samples from the exploration region, under an exploration probability ϵ, and an additional number of k_2 samples from the disclosed region. Let $F_{n+k_1+k_2}$ denote the empirical CDF estimated from these $n + k_1 + k_2$ non-IID samples. Then, for every l, m, n, k_1, k_2, and $\eta > 0$,*

$$\mathbb{P}\left[\sup_{\boldsymbol{x}\in\mathbb{R}^2}\left|F(\boldsymbol{x}) - F_{n+k_1+k_2}(\boldsymbol{x})\right| \geq \eta\right] \leq 4\exp\left(\frac{-2l(\eta-|\beta-\frac{l}{n}|)^2}{\min\left(\beta, \frac{l}{n}\right)^2}\right)$$
$$+ 4\exp\left(\frac{-2(m-l+k_1)\left(\eta-|\beta-\frac{l}{n}|-\left|\alpha-\beta-\frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\right|\right)^2}{\min\left(\alpha-\beta, \frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\right)^2}\right) + 4\exp\left(\frac{-2(n-m+k_2)\left(\eta-2\left|\alpha-\frac{l}{n}-\frac{n-l}{n}\frac{m-l+k_1}{n-l+k_1+\epsilon k_2}\right|\right)^2}{\min\left(1-\alpha, \frac{n-l}{n}\frac{n-m+\epsilon k_2}{n-l+k_1+\epsilon k_2}\right)^2}\right).$$

The proofs for Theorem 6 and 7 follow the same logic as the proofs for Theorem 2 and 3, with an application of the multivariate DKW inequality and the re-estimated α.

K Additional experiments

Additional experiments on CDF error bounds

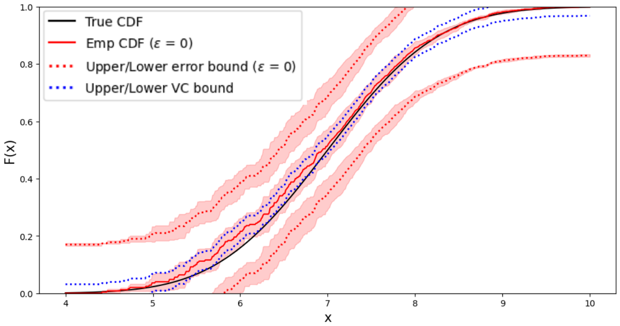

Figure 7: Comparison with our CDF error bound and existing VC bound

In this experiment, we aim to compare the performance of our bounds to the existing VC generalization error bound (Devroye et al., 2013, Cor. 12.2) through the confidence bounds ($\delta = 0.015$) around the empirical distribution function. We proceed with the same experiment setting on the synthetic data as in Fig 4, but without exploration. We run the experiments 5 times and report the average results with its error bars.

From Figure 7, we can clearly see that the existing VC generalization bound (blue dotted lines) fails to capture the true distribution, especially in the censored and exploration regions, from where we do not collect samples due to the censored feedback. This is because the discrepancy between the empirical and true distributions is large in regions with fewer samples, and the existing VC generalization bound does not account for censored feedback, decreasing the error bounds faster.