Towards Fair Disentangled Online Learning for Changing Environments

Chen Zhao*
charliezhaoyinpeng@gmail.com
Baylor University
Waco, Texas, USA

Feng Mi*
feng.mi@utdallas.edu
University of Texas at Dallas
Richardson, Texas, USA

Xintao Wu
xintaowu@uark.edu
University of Arkansas
Fayetteville, Arkansas, USA

Kai Jiang
kai.jiang@utdallas.edu
University of Texas at Dallas
Richardson, Texas, USA

Latifur Khan
lkhan@utdallas.edu
University of Texas at Dallas
Richardson, Texas, USA

Christan Grant
christan@ufl.edu
University of Florida
Gainesville, Florida, USA

Feng Chen
feng.chen@utdallas.edu
University of Texas at Dallas
Richardson, Texas, USA

ABSTRACT

In the problem of online learning for changing environments, data are sequentially received one after another over time, and their distribution assumptions may vary frequently. Although existing methods demonstrate the effectiveness of their learning algorithms by providing a tight bound on either dynamic regret or adaptive regret, most of them completely ignore learning with model fairness, defined as the statistical parity across different sub-population (*e.g.,* race and gender). Another drawback is that when adapting to a new environment, an online learner needs to update model parameters with a global change, which is costly and inefficient. Inspired by the sparse mechanism shift hypothesis [22], we claim that changing environments in online learning can be attributed to partial changes in learned parameters that are specific to environments and the rest remain invariant to changing environments. To this end, in this paper, we propose a novel algorithm under the assumption that data collected at each time can be disentangled with two representations, an environment-invariant semantic factor and an environment-specific variation factor. The semantic factor is further used for fair prediction under a group fairness constraint. To evaluate the sequence of model parameters generated by the learner, a novel regret is proposed in which it takes a mixed form of dynamic and static regret metrics followed by a fairness-aware long-term constraint. The detailed analysis provides theoretical guarantees for loss regret and violation of cumulative fairness constraints. Empirical evaluations on real-world datasets demonstrate our proposed method sequentially outperforms baseline methods in model accuracy and fairness.

CCS CONCEPTS

• **Computing methodologies** → **Artificial intelligence**; **Machine learning**; • **Applied computing** → *Law, social and behavioral sciences*; • **Social and professional topics** → User characteristics.

KEYWORDS

fairness, online learning, disentanglement, changing environments

ACM Reference Format:
Chen Zhao, Feng Mi, Xintao Wu, Kai Jiang, Latifur Khan, Christan Grant, and Feng Chen. 2023. Towards Fair Disentangled Online Learning for Changing Environments. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23), August 6–10, 2023, Long Beach, CA, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3580305.3599523

1 INTRODUCTION

Unlike offline learning approaches, where data is accumulated over time and collected at once, online learning assumes data batches are acquired as a continuous flow and sequentially received one after another, making it ideal for the real world. Although online learners can learn from new information in real-time as it arrives, state-of-the-art online learning algorithms may fail catastrophically when learning environments are dynamic and change over time, where changing environments refer to shifted distributions of data features between batches. Therefore, it requires online learning algorithms to adapt dynamically to new patterns in data sequences.

To address changing environments, adaptive regret [3] and dynamic regret [38] are introduced. Adaptive regret evaluates the learner's performance on any contiguous time intervals, and it is defined as the maximum static regret [38] over these intervals [3]. In contrast, dynamic regret handles changing environments from the perspective of the entire learning process. It allows the comparator changes over time. However, minimizing dynamic regret may be

*Both authors contributed equally to this research.

KDD '23, August 6–10, 2023, Long Beach, CA, USA

ACM ISBN 979-8-4007-0103-0/23/08…$15.00
https://doi.org/10.1145/3580305.3599523

less efficient because the learner needs to update model parameters with a global change against changing environments. Inspired by the sparse mechanism shift hypothesis [22], we state changing environments in online learning can be attributed to partial changes of parameters in a long run that are specific to environments. This implies that some parameters remain semantically invariant across different environments.

Existing fairness-aware online algorithms are developed with a focus on either static or adaptive regret. Learning fairness with dynamic regret for changing environments is barely touched. Data containing bias on some sensitive characters (*e.g.* race and gender) are likely collected sequentially over time. Group fairness is defined by the equality of a predictive utility across different data sub-populations, and predictions of a model are statistically independent on sensitive information. To control bias sequentially, the summation of fair constraints over time added to static loss regret is minimized [18]. It ensures the total violation of fair constraints sublinearly increases in time. Although the adaptive fair regret proposed in [37] is initially designed for online changing environments, it allows the learner to make decisions at some time that do not belong to the fair domain and assumes the total number of times is known in advance. Therefore, designing fairness-aware online algorithms associated with dynamic regret for changing environments becomes desirable.

In this paper, to address the problem of fairness-aware online learning, where a sequence of data batches (*e.g.* tasks) are collected one after another over time with changing task environments (see Fig. 1), we propose a novel regret metric, namely FairSDR, followed by long-term fairness-aware constraints. To adapt to dynamic environments, we state that shifts in data distributions can be attributed to partial updates in model parameters in a long run, with some remaining invariant to changing environments. Inspired by dynamic and static regret metrics, FairSDR and the violation of cumulative fair constraints are minimized and bounded with $O(\sqrt{T(1+P_T)})$ and $O(\sqrt{T})$, respectively, where T is the number of iterations and P_T is the path-length of the comparator sequence. To learn a sequence of model parameters satisfying the regret, we propose a novel online learning algorithm, namely FairDolce. In this algorithm, two learning networks are introduced, the representation learning network (RLN) and the prediction learning network (PLN). RLN disentangles an input with environment-invariant and environment-specific representations. It aims to ensure the semantic invariance of the learned presentation from RLN to all possible environments. Furthermore, the environment-invariant representations are used to predict class labels constrained with controllable fair notions in PLN. The main contributions of this paper are summarized[1]:

- We propose a novel regret FairSDR that compares the cumulative loss of the learner against any sequence of comparators for changing environments, where only partial parameters need to be adapted to the changed environments in the long run. The proposed new regret takes a mixed form of static and dynamic regret metrics, subject to a long-term fairness constraint.
- To adapt to changing environments, we postulate that model parameters are updated with a local change. An effective algorithm FairDolce is introduced, consisting of two networks: a representation learning network (RLN) and a prediction learning network (PLN). In RLN, datapoints are disentangled into two representations. With semantic representations, PLN is optimized under fair constraints.
- Theoretically grounded analysis justifies the effectiveness of the proposed method by demonstrating upper bounds $O(\sqrt{T(1+P_T)})$ for loss regret and $O(\sqrt{T})$ for violation of cumulative fair constraints.
- We validate the performance of our approach with state-of-the-art techniques on real-world datasets. Our results demonstrate FairDolce can effectively adapt both accuracy and fairness in changing environments and it shows substantial improvements over the best prior works.

[1]Code repository: https://github.com/harderbetter/fairdolce

2 RELATED WORK

Fairness-aware online learning. To sequentially ensure fairness guarantees at each time, a fairness-aware regret [20] considering the trade-off between model accuracy and fairness is devised and it provides a fairness guarantee held uniformly over time. Another trend [13, 18, 28, 36, 37] addressing this problem is to develop a new metric by adding a long-term fair constraint directly to the loss regret. However, when handling constrained optimization problems, the computational burden of the projection onto the fair domain may be too high when constraints are complex. For this reason, [18] relaxes the output through a simpler closed-form projection. Thereafter, a number of variants of [18] are proposed with theoretical guarantees by modifying stepsizes in [18] to an adaptive version, adjusting to stochastic constraints [27], and clipping constraints into a non-negative orthant [28]. Although such techniques achieve state-of-the-art bounds for static regrets and violation of fair constraints, they assume datapoints sampled at each time from a stationary distribution and make heavy use of the *i.i.d* assumption. This does not hold when the environment changes.

Online learning for changing environments. Because low static regret does not imply a good performance in changing environments, two regret metrics, dynamic regret [38] and adaptive regret [11], are devised to measure the learner's performance in changing environments. Adaptive regret handles changing environments from a local perspective by focusing on comparators in short intervals, in which geometric covering intervals [3, 14, 31] and data streaming techniques [8] are developed. CBCE [14] improved the strongly adapted regret bound by combing the sleeping bandits idea with the Coin Betting algorithm. AOD [31] targets both dynamic and adaptive regret and proposes theoretic guarantees to minimize both regrets simultaneously. Although existing methods achieve state-of-the-art performance, a major drawback is that they immerse in minimizing objective functions but ignore the model fairness of prediction. As the first work addressing the problem of online fairness learning for changing environments, FairSAOML [37] combines tasks with a number of sets with different lengths and develops an effective algorithm inspired by expert-tracking techniques. A major drawback of FairSAOML is that (1) it assumes some tasks are known in advance which leads to delays during the learning process; (2) by designing intervals with long lengths, it is hard for a learner to adapt to new environments without leaving

information from past environments behind. As a consequence, the adaptation of the learner to new environments may not perform well.

With concerns from existing works, to tackle the problem of fairness-aware online learning for changing environments, in this paper, we propose a novel regret and a learning algorithm, in which we assume only part of the model parameters is responsible for adapting to new environments and the rest are environment-invariant corresponding to fair predictions. Inspired by invariant learning strategies, the proposed algorithm FairDolce is used to accommodate changing environments and adaptively learn the model with accuracy and fairness.

3 PRELIMINARIES

Vectors are denoted by lowercase boldface letters. Scalars are denoted by lowercase italic letters. Sets are denoted by uppercase calligraphic letters. For more details refer to Appendix A.1.

3.1 Online Learning

In online learning, data batches $\mathcal{D}_t$, defined as tasks, arrive one after another over time. An online machine learner can learn from new information in real-time as they arrive. Specifically, at each time, the learner faces a loss function $f_t : \mathbb{R}^d \times \Theta \to \mathbb{R}$ which does not need to be drawn from a fixed distribution and could even be chosen adversarially over time [6]. The goal of the learner over all times T is to decide a sequence of model parameters $\{\boldsymbol{\theta}_t\}_{t=1}^T$ by an online learning algorithm, *e.g.,* follow the leader [9], that performs well on the loss sequence $\{f_t(\mathcal{D}_t, \boldsymbol{\theta}_t)\}_{t=1}^T$. Particularly, to evaluate the algorithm, a standard objective for online learning is to minimize some notion of regret, defined as the overall difference between the learner's loss $\sum_{t=1}^T f_t(\mathcal{D}_t, \boldsymbol{\theta}_t)$ and the best performance achievable by comparators.

Static regret. In general, one assumes that tasks collected over time are sampled from a fixed and stationary environment following the *i.i.d* assumption. Therefore, with a sequence of model parameters learned from the learner, the objective is to minimize the accumulative loss of the learned model to that of the best *fixed* comparator $\boldsymbol{\theta} \in \Theta$ in hindsight. This regret is typically referred to as static regret since the comparator is time-invariant.

$$R_s = \sum_{t=1}^T f_t(\mathcal{D}_t, \boldsymbol{\theta}_t) - \min_{\boldsymbol{\theta}\in\Theta} \sum_{t=1}^T f_t(\mathcal{D}_t, \boldsymbol{\theta}) \tag{1}$$

The goal of online learning under a stationary environment is to design algorithms such that static regret R_s sublinearly grows in T. However, low static regret does not necessarily imply a good performance in changing environment, where tasks are sampled from various distributions, since the time-invariant comparator $\boldsymbol{\theta}$ in Eq. (1) may behave badly. Tasks sequentially collected from non-stationary environments and distributions of them varying over time are more realistic. To address this limitation, recent advances [25, 33] have introduced enhanced regret metrics, *i.e.,* dynamic regret, to measure the learner's performance.

Dynamic regret. The dynamic regret [38] is defined as the difference between the cumulative loss of the learner and that of *a sequence* of comparators $\mathbf{u}_1, \cdots, \mathbf{u}_T \in \Theta$.

$$R_d = \sum_{t=1}^T f_t(\mathcal{D}_t, \boldsymbol{\theta}_t) - \sum_{t=1}^T f_t(\mathcal{D}_t, \mathbf{u}_t) \tag{2}$$

In fact, Eq. (2) is more general since it holds for any sequence of comparators and thus includes the static regret in Eq. (1). Therefore, minimizing dynamic regret can automatically adapt to the nature of environments, either stationary or dynamic. However, distinct from static regret, bounding dynamic regret is challenging because one needs to establish a universal guarantee that holds for any sequence of comparators [32]. An alternative solution for this challenge is to bound the regret in terms of some regularities of the comparator sequence, *e.g.,* path-length [38] defined in Eq. (9) which measures the temporal variability of the comparator sequence.

As alluded to in Sec. 1, most of the state-of-the-art online techniques ignore the significance of learning by being aware of model fairness, which is an important hallmark of human intelligence. To control bias, especially ensure group fairness across different sub-populations, cumulative fairness notions are considered as constraints added on regrets.

3.2 Group Fairness

In general, group fairness criteria used for evaluating and designing machine learning models focus on the relationships between the sensitive variables and the system output [24, 34, 35]. The problem of group unfairness prevention can be seen as a constrained optimization problem. For simplicity, we consider one binary sensitive label, *e.g.* gender, in this work. However, our ideas can be easily extended to many sensitive labels with multiple levels.

Let $\mathcal{P} = \mathcal{X} \times \mathcal{Z} \times \mathcal{Y} \times \mathcal{E}$ be the data space, where $\mathcal{X} \in \mathbb{R}^d$ is an input feature space, $\mathcal{Z} \in \{-1, 1\}$ is a sensitive space, $\mathcal{Y} \in \{0, 1\}$ is an output space for binary classification, and $\mathcal{E} \in \mathbb{N}$ denotes an environment space. Given a task $\mathcal{D} = \{(\mathbf{x}_i, z_i, y_i, e_i)\}_{i=1}^n \in \mathcal{P}$ in environment $e_i \in \mathcal{E}$, a fine-grained measurement to ensure group fairness in class label prediction is to design fair classifiers by controlling the notions of fairness between sensitive subgroups $\{z_i = 1\}_{i=1}^{n_1}$ and $\{z_i = -1\}_{i=1}^{n_{-1}}$ where $n_1 + n_{-1} = n$, *e.g.,* demographic parity [17, 24].

Definition 1 (Notions of Fairness [17, 24, 37]). *A classifier $\omega : \mathbb{R}^d \times \Theta \to \mathbb{R}$ is fair when its predictions are independent of the sensitive attribute $\mathbf{z} = \{z_i\}_{i=1}^n$. To get rid of the indicator function and relax the exact values, a linear approximated form of the difference between sensitive subgroups is defined [17],*

$$g(\mathcal{D}, \boldsymbol{\theta}) = \mathbb{E}_{(\mathbf{x},z,y,e)\sim\mathcal{P}}\Big[\frac{1}{\hat{p}_1(1-\hat{p}_1)}\Big(\frac{z+1}{2} - \hat{p}_1\Big)\omega(\mathbf{x}, \boldsymbol{\theta})\Big] \tag{3}$$

where $\hat{p}_1$ is an empirical estimate of pr_1. pr_1 is the proportion of samples in group $z = 1$ and correspondingly $1 - pr_1$ is the proportion of samples in group $z = -1$.

Notice that, in Eq. (3), when $\hat{p}_1 = \mathbb{P}_{(\mathbf{x},z,y,e)\in\mathcal{P}}(z = 1)$, the fairness notion $g(\mathcal{D}, \boldsymbol{\theta})$ is defined as the difference of demographic parity (DDP). Similarly, when $\hat{p}_1 = \mathbb{P}_{(\mathbf{x},z,y,e)\in\mathcal{P}}(y = 1, z = 1)$, $g(\mathcal{D}, \boldsymbol{\theta})$ is defined as the difference of equality of opportunity (DEO) [17]. Therefore, parameters $\boldsymbol{\theta}$ in the domain of a task are feasible if they strictly satisfy the fairness constraint $g(\mathcal{D}, \boldsymbol{\theta}) = 0$.

Motivations. To tackle the problem of fairness-aware online learning in changing environments, a learner needs to update model parameters with a global change, which is costly and inefficient. Inspired by the sparse mechanism shift hypothesis [22], we state changing environments in online learning can be attributed to

partial changes in learned parameters in the long run that are specific to environments. This implies that some parameters remain semantically invariant across different environments.

4 METHODOLOGY

4.1 Settings and Problem Formulation

We consider a general sequential setting where a learner is faced with tasks $\{\mathcal{D}_t\}_{t=1}^T$ one after another. Each of these tasks corresponds to a time, denoted as $t \in [T]$. At each time, the goal of the learner is to determine model parameters $\boldsymbol{\theta}_t$ using existing task pool $\{\mathcal{D}_i\}_{i=1}^{t-1}$ in a fair domain Θ that perform well for the task arrived at t. This is monitored by the loss function f_t and the fairness notion g_t, wherein the fair constraint $g_t(\mathcal{D}_t, \boldsymbol{\theta}_t) = 0$ is satisfied and $f_t(\mathcal{D}_t, \boldsymbol{\theta}_t)$ is minimized. To adapt to changing environments, crucially, model parameters $\boldsymbol{\theta}_t = \{\boldsymbol{\theta}_t^s, \boldsymbol{\theta}_t^v, \boldsymbol{\theta}_t^d, \boldsymbol{\theta}_t^{cls}\}$ can be partitioned into multiple elements, specifically in which $\boldsymbol{\theta}_t^s$ captures the semantic information of data through a semantic encoder $h_s : \mathcal{X} \times \Theta \to \mathcal{S}$, and $\boldsymbol{\theta}_t^{cls}$ is used for prediction under fair constraints. $\boldsymbol{\theta}_t^v$ and $\boldsymbol{\theta}_t^d$ are parameters, later introduced in Secs. 4.2 and 4.3, for encoding the environmental information and decoding latent representations, respectively, in order to adaptively train a good $\boldsymbol{\theta}_t^s$. For data batches sampled from heterogeneous distributions at different times, $\boldsymbol{\theta}_t^s$ corresponds to adapting to changing environments by encoding samples to a latent semantic space. With latent factors (representations) encoded from the semantic space as inputs, $\boldsymbol{\theta}_t^{cls}$ is time-invariant in the long run. The overall protocol for this setting is as follows:

(1) The learner selects semantic parameters $\boldsymbol{\theta}_t^s$ and classification parameters $\boldsymbol{\theta}_t^{cls}$ in the fair domain Θ.
(2) The world reveals a loss and fairness notion f_t and g_t.
(3) The learner incurs an instantaneous loss $f_t(h_s(\mathcal{D}_t, \boldsymbol{\theta}_t^s), \boldsymbol{\theta}_t^{cls})$ and fairness estimation $g(h_s(\mathcal{D}_t, \boldsymbol{\theta}_t^s), \boldsymbol{\theta}_t^{cls})$.
(4) Advance to the next time.

As mentioned in Sec. 3.1, the goal of the learner is to minimize regret under long-term constraints [18], defined as the summation of fair constraints over time. Since $\boldsymbol{\theta}_t^s$ adapts to different environments to encode semantic information from a latent invariant space and $\boldsymbol{\theta}_t^{cls}$ further takes semantic inputs for fair prediction, let $\{\boldsymbol{\theta}_t^s, \boldsymbol{\theta}_t^{cls}\}_{t=1}^T$ be the sequence of parameters generated at the *Step* (1) of the protocol. We propose *a novel fairness-aware regret for changing environments*, namely FairSDR, defined as

$$FairSDR = \sum_{t=1}^{T} f_t(h_s(\mathcal{D}_t, \boldsymbol{\theta}_t^s), \boldsymbol{\theta}_t^{cls}) - \min_{\boldsymbol{\theta}^{cls} \in \Theta} \sum_{t=1}^{T} f_t(h_s(\mathcal{D}_t, \mathbf{u}_t^s), \boldsymbol{\theta}^{cls})$$

$$\text{subject to} \quad \sum_{t=1}^{T} \left\| \left[g(h_s(\mathcal{D}_t, \boldsymbol{\theta}_t^s), \boldsymbol{\theta}_t^{cls}) \right]_+ \right\| = 0$$

where $[\cdot]_+$ is the projection onto the non-negative space. Similar to $\{\mathbf{u}_1, \cdots, \mathbf{u}_T\}$ denoted in Eq. (2), $\{\mathbf{u}_1^s, \cdots, \mathbf{u}_T^s\}$ are a sequece of semantic comparators to $\{\boldsymbol{\theta}_1^s, \cdots, \boldsymbol{\theta}_T^s\}$, where each corresponds to an underlying environment. $\boldsymbol{\theta}^{cls}$ is the best-fixed comparator for fair classification, which is time-invariant.

Remarks. In contrast to the regret proposed in [37] in which it is extended from the interval-based strongly adaptive regret, and it aims to minimize the maximum static regret for all intervals on the undivided model parameter, FairSDR takes the mixed form of static and dynamic regrets. Furthermore, [37] employs the meta-learning framework in which the function inside f_t is designed for interval-level learning with gradient steps on $\boldsymbol{\theta}$. However, in FairSDR, h_s encodes an input to a semantic representation through a neural network on $\boldsymbol{\theta}^s$, which is part of $\boldsymbol{\theta}$.

4.2 Assumptions for Invariance

Recall that in the learning protocol mentioned in Sec. 4.1, the main goal for the learner is to generate the parameter sequence $\{\boldsymbol{\theta}_t^s, \boldsymbol{\theta}_t^{cls}\}_{t=1}^T$ in Sec. 4.1 that performs well on the loss sequence and the long-term fair constraints. We make the following assumptions.

Assumption 1 (Shared Semantic Space). *Given a task $\{(\mathbf{x}_i, z_i, y_i, e_i)\}_{i=1}^n$ sampled from a particular environment $e_i \in \mathcal{E}$, we assume that each datapoint in the task is generated from*

- *a semantic factor $\mathbf{s}_i = h_s(\mathbf{x}_i, \boldsymbol{\theta}^s) \in \mathcal{S}$, where $\mathcal{S}$ refers to a semantic space shared by all environment $\mathcal{E}$;*
- *a variation factor $\mathbf{v}_i = h_v(\mathbf{x}_i, \boldsymbol{\theta}^v) \in \mathcal{V}$ where $\mathbf{v}_i$ is specific to the individual environment e_i.*

where $h_v : \mathcal{X} \times \Theta \to \mathcal{V}$ is a variation encoder parameterized by $\boldsymbol{\theta}^v$. We assume that each environment e_i is represented by specific variation factor $h_v(\mathbf{x}_i, \boldsymbol{\theta}^v)$.

This assumption is closely related to the shared latent space assumption in [16], wherein [16] assumes a fully shared latent space. We postulate that only the semantic space can be shared across environments whereas the variation factor is environment specific, which is a more reasonable assumption when the cross-environment mapping is many-to-many. In other words, given datapoints in various environments, each can be encoded into semantic and variation factors within the same semantic space but with different variation factors depending on the environments.

Under Assumption 1, each datapoint is able to be disentangled with semantic and variation factors. With two datapoints sampled from the same environment e_i, given a decoder $D : \mathcal{S} \times \mathcal{V} \times \Theta \to \mathcal{X}$, we assume that

Assumption 2 (Data Invariance under Homogeneous Environments). *Given a semantic encoder h_s, a variation encoder h_v, and a decoder D, for any $\mathbf{x}_i, \mathbf{x}_j \in \mathcal{X}, i \neq j$ sampled in the same environment $e \in \mathcal{E}$, it holds $\mathbf{x}_i = D(h_s(\mathbf{x}_i, \boldsymbol{\theta}^s), h_v(\mathbf{x}_j, \boldsymbol{\theta}^v), \boldsymbol{\theta}^d)$.*

Assumption 2 enforces the data invariance of the original input $\mathbf{x}_i$ and the one that $D(h_s(\mathbf{x}_i, \boldsymbol{\theta}^s), h_v(\mathbf{x}_j, \boldsymbol{\theta}^v), \boldsymbol{\theta}^d)$ reconstructs jointly from semantic and variation latent factors when the latter remains but the former varies.

Assumption 3 (Class Invariance under Heterogeneous Environments [30]). *We assume that inter-environment variation is solely characterized by the environment shift in the distribution $\mathbb{P}(X, E)$. As a consequence, we assume that $\mathbb{P}(Y|X, E)$ is stable across environments. Similar to [21, 30], given two datapoints $(\mathbf{x}_i, z_i, y, e_i)$ and $(\mathbf{x}_j, z_j, y, e_j)$, we assume the following holds*

$$\mathbb{P}(Y = y|X = \mathbf{x}_i, E = e_i) = \mathbb{P}(Y = y|(X = D(h_s(\mathbf{x}_i, \boldsymbol{\theta}^s), h_v(\mathbf{x}_j, \boldsymbol{\theta}^v), \boldsymbol{\theta}^d), E = e_j),$$
$$\forall \mathbf{x}_i, \mathbf{x}_j \in \mathcal{X}, e_i, e_j \in \mathcal{E}, i \neq j$$

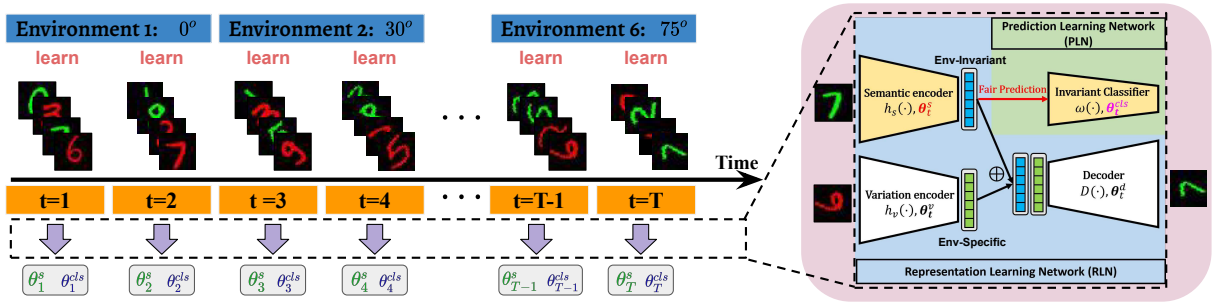

Figure 1: A graphical illustration of the proposed framework using Rotated-Colored-MNIST dataset. (Left) Each angle within $\{0, 15, 30, 45, 60, 75\}$ represents an environment. In the problem of fairness-aware online learning for changing environments, data batches arrive one after another over time. Parameters sequence $\{\theta_t^s, \theta_t^{cls}\}_{t=1}^T$ are learned through the proposed model on the right. (Right) The model consists of two learning networks, RLN and PLN. The semantic and variation encoders disentangle an input with two factors (representations). Under Assumptions 2 and 3, the decoder takes both factors and generates new data by diversifying the variation across environments. Semantic factors go through the classifier presented in PLN under fair constraints and further output fair predictions. We claim that when T is large enough, only a subset of the parameters sequence, $\{\theta_t^s\}_{t=1}^T$, are updated to adapt to changing environments.

This assumption shows that the prediction depends only on the semantic factor $h_s(\mathbf{x}, \theta^s)$ regardless of the variation factor $h_v(\mathbf{x}, \theta^v)$. Furthermore, the semantic factors are used for fair prediction under fairness constraints.

4.3 Learning Dynamically for Adaptation

As the motivation stated in Sec. 3, an efficient online algorithm is expected to partially update model parameters (*i.e.,* θ_t^s) to adapt to changing environments sequentially and to remain the rest (*i.e.,* θ_t^{cls}). As the illustration shown in Fig. 1, a novel online framework for changing environments is proposed with two separate networks. The representation learning network (RLN) aims to learn a good semantic encoder h_s that is able to accurately disentangle semantic representations within various environments, associated with the variation encoder h_v and the decoder D. The prediction learning network (PLN) solely consists of the classifier ω and it takes semantic representations from RLN and outputs fair predictions under fair constraints, which is invariant to environments.

Specifically in RLN, to learn a good semantic encoder h_s, at each time t we consider a data batch $Q_t = \{(\mathbf{r}_{1,q,t}, \mathbf{r}_{2,q,t}, \mathbf{r}_{3,q,t}, \mathbf{r}_{4,q,t})\}_{q=1,t}^{Q}$ containing multiple quartet data pairs sampled from existing task pool $\{\mathcal{D}_i\}_{i=1}^{t-1}$, where Q denotes the number of quartet pairs in $|Q_t|$.

- $\mathbf{r}_{1,q,t} = (\mathbf{x}_{a,t}, z_{a,t}, y_t, e_t)$ with class y_t and environment e_t
- $\mathbf{r}_{2,q,t} = (\mathbf{x}_{b,t}, z_{b,t}, y'_t, e_t)$ with class y'_t and environment e_t
- $\mathbf{r}_{3,q,t} = (\mathbf{x}_{c,t}, z_{c,t}, y_t, e'_t)$ with class y_t and environment e'_t
- $\mathbf{r}_{4,q,t} = (\mathbf{x}_{d,t}, z_{d,t}, y'_t, e'_t)$ with class y'_t and environment e'_t

Notice that $\mathbf{r}_{1,q,t}$ and $\mathbf{r}_{2,q,t}$ (same to $\mathbf{r}_{3,q,t}$ and $\mathbf{r}_{4,q,t}$) share the same environment label e_t but different labels y_t and y'_t. $\mathbf{r}_{1,q,t}$ and $\mathbf{r}_{3,q,t}$ (same to $\mathbf{r}_{2,q,t}$ and $\mathbf{r}_{4,q,t}$) share the same label y_t but different environments e_t and e'_t. We view $\mathbf{r}_{3,q,t}$ ($\mathbf{r}_{4,q,t}$) is an alternative pair to $\mathbf{r}_{1,q,t}$ ($\mathbf{r}_{2,q,t}$) with changing environments. For simplicity, we omit the subscripts q and t.

Under Assumption 2, for $(\mathbf{r}_1, \mathbf{r}_2)$ and $(\mathbf{r}_3, \mathbf{r}_4)$ within the *same environment but different labels*, the data reconstruction loss $\mathcal{L}_{recon}$ is given:

$$\mathcal{L}_{recon}^q = dist[\mathbf{x}_a, D(\mathbf{s}_a, \mathbf{v}_b, \theta_t^d)] + dist[\mathbf{x}_c, D(\mathbf{s}_c, \mathbf{v}_d, \theta_t^d)] \quad (4)$$

where $\mathbf{s}_a = h_s(\mathbf{x}_a, \theta_t^s)$, $\mathbf{s}_c = h_s(\mathbf{x}_c, \theta_t^s)$, $\mathbf{v}_b = h_v(\mathbf{x}_b, \theta_t^v)$, and $\mathbf{v}_d = h_v(\mathbf{x}_d, \theta_t^v)$. $dist : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ indicates a distance metric, where we use ℓ_1 norm in the experiments.

Similarly, under Assumption 3, for $(\mathbf{r}_1, \mathbf{r}_3)$ and $(\mathbf{r}_2, \mathbf{r}_4)$ with the *same label but different environments*, the class invariance loss $\mathcal{L}_{inv}$ is given:

$$\mathcal{L}_{inv}^q = \ell_{CE}\Big(\omega(h_s(\mathbf{x}_{a\to c}, \theta_t^s), \theta_t^{cls}), y\Big) + \ell_{CE}\Big(\omega(h_s(\mathbf{x}_{b\to d}, \theta_t^s), \theta_t^{cls}), y'\Big) \quad (5)$$

where $\mathbf{x}_{a\to c} = D(\mathbf{s}_a, \mathbf{v}_c, \theta_t^d)$, $\mathbf{x}_{b\to d} = D(\mathbf{s}_b, \mathbf{v}_d, \theta_t^d)$, and $\ell_{CE} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the cross-entropy loss function.

Finally, to ensure prediction accuracy within a fair domain, we combine $(\mathbf{r}_{1,q}, \mathbf{r}_{2,q}, \mathbf{r}_{3,q}, \mathbf{r}_{4,q})$ together over the batch Q_t to estimate $\mathcal{L}_{cls}$ and $\mathcal{L}_{fair}$:

$$\begin{aligned} \mathcal{L}_{cls} &= f_t(Q_t, \theta_t^s \oplus \theta_t^{cls}) \\ &= \sum_{q=1}^{Q} \sum_{k}^{\{a,b,c,d\}} \ell_{CE}\big(\omega(h_s(\mathbf{x}_{k,q}, \theta_t^s), \theta_t^{cls}), y_{k,q}\big) \\ \mathcal{L}_{fair} &= \sum_{q=1}^{Q} g(Q_t, \theta_t^s \oplus \theta_t^{cls}) \end{aligned} \quad (6)$$

where $\oplus$ denotes concatenation operator between θ_t^s and θ_t^{cls}.

4.4 A Practical Online Algorithm: FairDolce

In practice, requirements for remaining data invariance for datapoints sampled in the same environment with different labels and for keeping class invariance for datapoints sampled with the same label within various environments are hard to be satisfied. Similar to the fairness constraint, it is a strict equality constraint that is difficult to enforce in practice. To alleviate some of such difficulties, we relax the loss functions with empirical constants that

$$\begin{aligned} &\mathcal{L}_{fair} \le \epsilon_1 \\ &\mathcal{L}_{recon} = \frac{1}{Q}\sum_{q=1}^{Q} \mathcal{L}_{recon}^q \le \epsilon_2; \quad \mathcal{L}_{inv} = \frac{1}{Q}\sum_{q=1}^{Q} \mathcal{L}_{inv}^q \le \epsilon_3 \end{aligned} \quad (7)$$

$\epsilon_1, \epsilon_2, \epsilon_3 > 0$ are fixed margins that control the extent to violations.

$$\mathcal{L}_{total} = \mathcal{L}_{cls} + \lambda_{t,1}(\mathcal{L}_{fair} - \epsilon_1) + \lambda_{t,2}(\mathcal{L}_{recon} - \epsilon_2) + \lambda_{t,3}(\mathcal{L}_{inv} - \epsilon_3) \quad (8)$$

Furthermore, we propose a primal-dual Algorithm 1 for efficient optimization, wherein it alternates between optimizing

Algorithm 1 FairDolce

1: Input: batch size Q, learning rate η_1, η_2, margin $\epsilon_1, \epsilon_2, \epsilon_3$.
2: Randomly initialize $\boldsymbol{\theta}_{t=0} \in \Theta$ and $\lambda_{0,1}, \lambda_{0,2}, \lambda_{0,3} \in \mathbb{R}_+$
3: Initial the domain buffer as empty, $\mathcal{U} \leftarrow [\,]$.
4: Initial the task buffer as empty, $\mathcal{T} \leftarrow [\,]$.
5: **for** each $t \in [T]$ **do**
6: Record the performance of $(\boldsymbol{\theta}^s_{t-1}, \boldsymbol{\theta}^{cls}_{t-1})$ on $\mathcal{D}_t$.
7: **if** $e_t \notin \mathcal{U}$ **then**
8: $\mathcal{U} \leftarrow \mathcal{U} \cup \{e_t\}$
9: **end if**
10: Assign $\boldsymbol{\theta}_t \leftarrow \boldsymbol{\theta}_{t-1}$, $\lambda_{t,1} \leftarrow \lambda_{t-1,1}$, $\lambda_{t,2} \leftarrow \lambda_{t-1,2}$, $\lambda_{t,3} \leftarrow \lambda_{t-1,3}$
11: **for** $n = 1, 2 \cdots$ steps **do**
12: **if** $|\mathcal{U}| \neq 1$ **then**
13: Randomly sample a batch $\mathcal{Q}_t \subset \mathcal{T}$ indicated in Sec. 4.3.
14: Compute $\mathcal{L}^q_{recon}$ and $\mathcal{L}^q_{inv}$ using Eq. (4) and (5) for each quartet pair.
15: $\mathcal{L}_{recon} = \frac{1}{Q}\sum_{q=1}^{Q} \mathcal{L}^q_{recon}$ and $\mathcal{L}_{inv} = \frac{1}{Q}\sum_{q=1}^{Q} \mathcal{L}^q_{inv}$
16: **else**
17: Randomly sample a batch of doublet data pairs $\mathcal{Q}_t = \{((\mathbf{x}_{i,q,t}, z_{i,q,t}, y_{i,q,t}, e), (\mathbf{x}_{j,q,t}, z_{j,q,t}, y_{j,q,t}, e))\}_{q=1}^{Q}$, where $\mathcal{Q}_t \subset \mathcal{T}$.
18: Compute $\mathcal{L}^q_{recon}$ using Eq. (4) for each doublet pair.
19: $\mathcal{L}_{recon} = \frac{1}{Q}\sum_{q=1}^{Q} \mathcal{L}^q_{recon}$
20: Set $\mathcal{L}_{inv} = 0$
21: **end if**
22: Compute $\mathcal{L}_{cls}$, $\mathcal{L}_{fair}$ using Eq. (6).
23: Compute $\mathcal{L}_{total}$ using Eq. (8).
24: $\boldsymbol{\theta}^s_t \leftarrow \text{Adam}(\mathcal{L}_{total}, \boldsymbol{\theta}^s_t, \eta_1)$
25: $\boldsymbol{\theta}^v_t \leftarrow \text{Adam}(\lambda_{t,2} \cdot \mathcal{L}_{recon} + \lambda_{t,3} \cdot \mathcal{L}_{inv}, \boldsymbol{\theta}^v_t, \eta_1)$
26: $\boldsymbol{\theta}^d_t \leftarrow \text{Adam}(\lambda_{t,2} \cdot \mathcal{L}_{recon} + \lambda_{t,3} \cdot \mathcal{L}_{inv}, \boldsymbol{\theta}^d_t, \eta_1)$
27: $\boldsymbol{\theta}^{cls}_t \leftarrow \text{Adam}(\mathcal{L}_{cls} + \lambda_{t,1} \cdot \mathcal{L}_{fair} + \lambda_{t,3} \cdot \mathcal{L}_{inv}, \boldsymbol{\theta}^{cls}_t, \eta_1)$
28: $\lambda_{t,1} \leftarrow \max\{\lambda_{t,1} + \eta_2 \cdot (\mathcal{L}_{fair} - \epsilon_1), 0\}$
29: $\lambda_{t,2} \leftarrow \max\{\lambda_{t,2} + \eta_2 \cdot (\mathcal{L}_{recon} - \epsilon_2), 0\}$
30: **if** $|\mathcal{U}| \neq 1$ **then**
31: $\lambda_{t,3} \leftarrow \max\{\lambda_{t,3} + \eta_2 \cdot (\mathcal{L}_{inv} - \epsilon_3), 0\}$
32: **end if**
33: **end for**
34: $\mathcal{T} \leftarrow \mathcal{T} \cup \{\mathcal{D}_t\}$
35: **end for**

$\boldsymbol{\theta}_t = \{\boldsymbol{\theta}^s_t, \boldsymbol{\theta}^v_t, \boldsymbol{\theta}^d_t, \boldsymbol{\theta}^{cls}_t\}$ at each time via minimizing the empirical Lagrangian with fixed dual $\lambda_t = \{\lambda_{t,1}, \lambda_{t,2}, \lambda_{t,3}\}$ corresponding for $\mathcal{L}_{fair}$, $\mathcal{L}_{recon}$ as well as $\mathcal{L}_{inv}$ and updating the dual variable according to the minimizer (lines 24-32). The primal-dual iteration has clear advantages over stochastic gradient descent in solving constrained optimization problems. Specifically, it avoids introducing extra balancing hyperparameters. Moreover, it provides convergence guarantees once we have sufficient iterations and a sufficiently small step size [30].

Moreover, because each task corresponds to a timestamp t and an unknown environment before $\mathcal{D}_t$ arrives, the collected task pool $\{\mathcal{D}_i\}_{i=1}^{t-1}$ may be sampled from a single environment. In this sense, instead of using a batch stated in Sec. 4.3 with multiple quartet pairs, a sampled batch with doublet pairs containing $\{(\mathbf{r}_{1,q,t}, \mathbf{r}_{1,q,t})\}_{q=1,t}^{Q}$ is considered. As a consequence, the class invariance loss in Eq. (5) is set to zero (lines 17-20).

5 ANALYSIS

We first state assumptions about the online learning problem for changing environments that are largely used in [6, 31, 32, 37]. Then we provide theoretical guarantees for the proposed FairSDR regarding the loss regret and violation of cumulative fair constraints.

Assumption 4 (Bounded Parameter Domain). *The parameter domain Θ has a bounded diameter D and contains the origin.*

$$\max_{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta} ||\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2|| \leq D, \quad \forall \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta$$

Assumption 5 (Convexity). *Domain Θ is convex and closed. The loss function f_t and the fair function g are convex.*

Assumption 6 (F–Lipschitz). *There exists a positive constant F such that*

$$\max_{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta} |f_t(\cdot, \boldsymbol{\theta}_1) - f_t(\cdot, \boldsymbol{\theta}_2)| \leq F,$$
$$\max_{\boldsymbol{\theta}_1 \in \Theta} ||g(\cdot, \boldsymbol{\theta}_1)|| \leq F, \quad \forall \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta, \forall t \in [T]$$

Assumption 7 (Bounded gradient). *The gradients $\nabla f_t(\boldsymbol{\theta})$ and $\nabla g(\boldsymbol{\theta})$ exist, and they are bounded by a positive constant G on Θ, i.e.,*

$$\max_{\boldsymbol{\theta} \in \Theta} ||\nabla f_t(\cdot, \boldsymbol{\theta})|| \leq G, \quad \max_{\boldsymbol{\theta} \in \Theta} ||\nabla g(\cdot, \boldsymbol{\theta})|| \leq G, \quad \forall \boldsymbol{\theta} \in \Theta, \forall t \in [T]$$

Examples where these assumptions hold include logistic regression and $L2$ regression over a bounded domain. As for constraints, a family of fairness notions, such as DDP stated in Eq. (3) of Sec. 3.2, are applicable as discussed in [17]. For simplicity, in this section, we omit $\mathcal{D}$ used in $f_t, \forall t$ and g.

As introduced in Sec. 4.1, a sequence of parameters $\{\boldsymbol{\theta}^s_1, \cdots, \boldsymbol{\theta}^s_T, \boldsymbol{\theta}^{cls}_1, \cdots, \boldsymbol{\theta}^{cls}_T\}$ generated by the learner are evaluated with comparator sequence $\{\mathbf{u}^s_1, \cdots, \mathbf{u}^s_T, \boldsymbol{\theta}^{cls}\}$ in FairSDR. We claim that FairSDR takes a mixed form of the static and dynamic regrets with respect to $\{\boldsymbol{\theta}^{cls}_t\}_{t=1}^T$ and $\{\boldsymbol{\theta}^s_t\}_{t=1}^T$, respectively. Since the comparator $\boldsymbol{\theta}^{cls}$ in the static regret is performed as the best fixed one in hindsight, intuitively the comparator sequence can be extended to $\{\mathbf{u}^s_1, \cdots, \mathbf{u}^s_T, \boldsymbol{\theta}^{cls}, \cdots, \boldsymbol{\theta}^{cls}\}$ by making T copies of $\boldsymbol{\theta}^{cls}$. For simplicity, we denote the sequence of the learner's parameters and comparators as $\{\boldsymbol{\theta}^l_t\}_{t=1}^T$ and $\{\mathbf{u}^c_t\}_{t=1}^T$, respectively, where $\boldsymbol{\theta}^l_t := \boldsymbol{\theta}^s_t \oplus \boldsymbol{\theta}^{cls}_t$ and $\mathbf{u}^c_t := \mathbf{u}^s_t \oplus \boldsymbol{\theta}^{cls}, \forall t \in [T]$.

Furthermore, different from the static regret introduced in Eq. (1), it is impossible to achieve a sub-linear upper bound using dynamic regret in general. Instead, we can bound the dynamic regret in terms of some certain regularity of the comparator sequence or the function sequence, such as the path-length [38] which measures the temporal variability of the comparator sequence.

$$P_T = \sum_{t=1}^{T} ||\mathbf{u}^c_{t+1} - \mathbf{u}^c_t||_2 \tag{9}$$

Finally, under Assumptions 4 to 7 and Eq. (9), we state the key Theorem 1 that the proposed FairSDR enjoys theoretic guarantees for both loss regret and violation of the long-term fairness constraint in the long run for Algorithm 1.

Table 1: Comparison of upper bounds in loss regret and constraint violations for changing environments across methods.

Algorithms	*M. Zinkevich* [38]	Ader [32]	AOD [31]	CBCE[14]	FairSAOML [37]	FairSDR (Ours)
Loss Regret	$O(T^{1/2}(1+P_T))$	$O((T(1+P_T))^{1/2})$	$O\big((\tau \log T)^{1/2}\big)$	$O\big((\tau \log T)^{1/2}\big)$	$O\big((\tau \log T)^{1/2}\big)$	$O((T(1+P_T))^{1/2})$
Constraint Violations	-	-	-	-	$O\big((\tau T \log T)^{1/4}\big)$	$O(T^{1/2})$

Theorem 1. *Suppose Assumptions 4 to 7 hold, let* $\{\boldsymbol{\theta}_t^s, \boldsymbol{\theta}_t^{cls}\}_{t=1}^T$ *be the sequence generated by the online learner in Algorithm 1 and* $\{\mathbf{u}_t^s\}_{t=1}^T \cup \{\boldsymbol{\theta}^{cls}\}$ *be the comparator sequence, setting adaptive learning rates with*

$$\eta_{1,t} = \eta_{1,0}/\sqrt{T}, \quad \eta_{2,t} = \eta_{2,0}/\sqrt{\eta_{1,t}}, \forall t \in [T]$$

where $\eta_{1,0} > 0$ *and* $\eta_{2,0} \in (0, \frac{1}{\sqrt{2}G})$ *are constants. We have*

$$\sum_{t=1}^{T} f_t(h_s(\boldsymbol{\theta}_t^s), \boldsymbol{\theta}_t^{cls}) - \min_{\boldsymbol{\theta}^{cls} \in \Theta} \sum_{t=1}^{T} f_t(h_s(\mathbf{u}_t^s), \boldsymbol{\theta}^{cls}) = O\Big(\sqrt{T(1+P_T)}\Big)$$

$$\sum_{t=1}^{T} \Big\| \big[g(h_s(\boldsymbol{\theta}_t^s), \boldsymbol{\theta}_t^{cls})\big]_+ \Big\| = O(\sqrt{T})$$

Proof. Proof of Theorem 1 is given in Appendix C. □

Discussion. Under Assumptions 4 to 7, we provide comparable bounds for FairSDR with respect to both loss regret and violation of fair constraints. Tab. 1 lists a number of state-of-the-art works focusing on the problem of online learning in changing environments, where ours are added at the end. AOD [31], CBCE [14], and FairSAOML [37] address this problem by proposing strongly adaptive regret. In contrast to dynamic regret, strongly adaptive regret handles changing environments from a local perspective by proposing a set of intervals ranging from τ tasks. Ader [32] and *M. Zinkevich* [38] tackle this problem using dynamic regret using the length-path regularity in Eq. (9). Although the loss regret we derived for FairSDR is comparable to the one in Ader, the latter ignores the long-term fair constraint which is essential for fair online learning.

6 EXPERIMENTAL SETTINGS

In previous sections, we derive a theoretically principled algorithm assuming convexity everywhere. However, it has been known that deep learning models provide advanced performance in real-world applications, but they have a non-convex landscape with challenging theoretical analysis. Taking inspiration from the success of deep learning, we empirically evaluate the proposed algorithm FairDolce using neural networks in this section.

Datasets. We consider four datasets: Rotated-Colored-MNIST (rcMNIST), New York Stop-and-Frisk [15], Chicago Crime [34], and German Credit [1] to evaluate our FairDolce against state-of-the-art baselines, where rcMNIST is an image data and the other three are tabular datasets. We include the visualization of rcMNIST in Fig. 2. **(1) Rotated-Colored-MNIST** is extended from the Rotated-MNIST dataset [7], which consists of 10,000 digits from 0 to 9 with different rotated angles where environments are determined by angles $\{0, 15, 30, 45, 60, 75\}$. For simplicity, we consider binary classification where digits are labeled with 0 and 1 for digits from 0-4 and 5-9, respectively. For fairness concerns, each image has a green or red digit color as the sensitive attribute. We intentionally make correlations between labels and digit colors for each corresponding environment ranging from $\{0.9, 0.7, 0.5, 0.3, 0.1, 0.05\}$. We further divide data from each environment equally into 3 subsets, where each is considered a task. For 6 environments, there is are total of 18 tasks and each arrives one after another over time in order. **(2) New York Stop-and-Frisk** [15] is a real-world dataset on policing in New York City in 2011. It documents whether a pedestrian who was stopped on suspicion of weapon possession would in fact possess a weapon. We consider race (*i.e.,* black and non-black) as the sensitive label for each datapoint. Since this data consists of data from 5 cities in New York City, Manhattan, Brooklyn, Queens, Bronx, and Staten, data collected from each city is considered as an individual environment. To adapt to the setting of online learning, data in each environment is further split into 3 tasks, 15 tasks in total, where each task corresponds to a month's set of data of a city. **(3) Chicago Crime** [34] dataset contains information including demographics information (*e.g.,* race, gender, age, population, *etc.*), household, education, unemployment status, *etc.* We use race (*i.e.,* black and non-black) as the sensitive label. It consists of 16 tasks and each corresponds to a county of Chicago city as an environment. This dataset is initially used for multi-task fair regression learning in [34], where crime counts are used as continuous labels for data records. In our experiments, we categorize crime counts into binary labels, high (≥ 6) and low (< 6). **(4) German Credit** [1] dataset contains 1000 datapoints with 20 features. Gender (*i.e.,* male and female) is used as sensitive attribute and credit risk (*i.e.,* good and bad) is the target. Following [23, 37], to generate dynamic environments, we construct a larger dataset by combining three copies of the original data and flipping the original values of non-sensitive attributes by multiplying -1 for the middle copy. Therefore, each copy is considered as an environment. Each data copy is split into 2 tasks by time and there are 6 tasks in total.

Evaluation Metrics. Three popular evaluation metrics to estimate fairness are used and each allows quantifying the extent to model bias.

- *Demographic Parity* (DP) [4] is formalized as

$$\text{DP} = \begin{cases} \mathbb{P}(\hat{Y}=1|Z=-1) \Big/ \mathbb{P}(\hat{Y}=1|Z=1), & \text{if DP} \leq 1 \\ \mathbb{P}(\hat{Y}=1|Z=1) \Big/ \mathbb{P}(\hat{Y}=1|Z=-1), & \text{otherwise} \end{cases}$$

  This is also known as a lack of disparate impact [5]. A value closer to 1 indicates fairness.
- *Equalized Odds* (EO) [10] is formalized as

$$\text{EO} = \begin{cases} \mathbb{P}(\hat{Y}=1|Z=-1,Y=1) \Big/ \mathbb{P}(\hat{Y}=1|Z=1,Y=1), & \text{if EO} \leq 1 \\ \mathbb{P}(\hat{Y}=1|Z=1,Y=1) \Big/ \mathbb{P}(\hat{Y}=1|Z=-1,Y=1), & \text{otherwise} \end{cases}$$

  EO requires that $\hat{Y}$ has equal true positive and false negative rates between subgroups $z=-1$ and $z=1$. Same to DP, a value closer to 1 indicates fairness.
- *Mean Difference* (MD) [29] is a form of statistical parity, applied to the classification decisions, measuring the difference in the

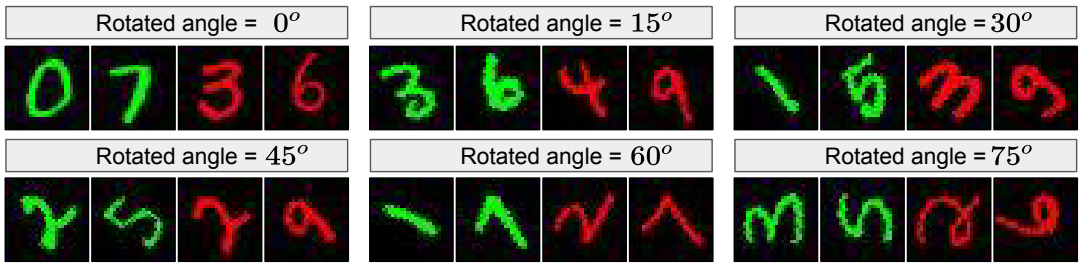

Figure 2: Visualization of the Rotated-Colored-MNIST dataset.

proportion of positive class of individuals in sub-groups.

$$\text{MD} = \left| \frac{\sum_{i:z_i=1} \hat{y}_i}{\sum_{i:z_i=1} 1} - \frac{\sum_{i:z_i=-1} \hat{y}_i}{\sum_{i:z_i=-1} 1} \right|$$

A value closer to 0 indicates fairness.

Baselines. We compare the performance of our proposed FairDolce with six baseline methods from three perspectives: online learning for changing environments (AOD [31], CBCE [14]), online fairness learning (FairFML [36], FairAOGD [13], FairGLC [28],), and the state-of-the-art online fairness learning for changing environments (FairSAOML [37]). AOD minimizes the strongly adaptive regret by running multiple online gradient descent algorithms over a set of dense geometric covering intervals. CBCE adapts changing environment in an online learning paradigm by combining the idea of sleeping bandits with the coin betting algorithm. FairFML controls bias in an online working paradigm and aims to attain zero-shot generalization with task-specific adaptation. FairFML focuses on a static environment and assumes tasks are sampled from an unchangeable distribution. FairAOGD is proposed for online learning with long-term constraints. In order to fit bias-prevention and compare them to FairDolce, we specify such constraints as DDP stated in Eq. (3). FairGLC rectifies FairAOGD by square-clipping the constraints in place of $g_i(\cdot), \forall i$. FairSAOML addresses fair online learning in changing environments by dynamically activating a subset of learning processes at each time through different combinations of task sets.

Architectures. For the rcMNIST image dataset, all images are resized to 28×28. Following [30], the semantic encoder and the style encoder consist of 4 strided convolutional layers followed by ReLU activation functions and Batch Normalization [12]. The decoder consists of 1 upsampling layer and 6 strided convolutional layers activated by ReLU. The classifier is performed by 2 FC hidden layers activated by ReLU. For tabular datasets (*i.e.,* New York Stop-and-Frisk, Chicago Crime, and German Credit), followed by [19], both encoders and the decoder contain one FC layer followed by LeakyReLU activation functions. The network architecture for the classifier is 1 FC layer activated by Sigmoid. Details for hyperparameters tuning are provided in Appendix A.2.

7 RESULTS

7.1 Adaptability for Changing Environments

As shown in the first three columns in Fig. 3, model performance is sequentially evaluated by fairness metrics (*i.e.,* DP, EO, and MD) introduced in Sec. 6. Our results demonstrate FairDolce outperforms baseline methods by giving the highest DP and EO values and the lowest MD overall. Specifically, it eventually meets the fair criteria of "80%-rule" [2] where DP and EO at the last several times are beyond 0.8. The last column of Fig. 3 shows the change of model accuracies over time. We claim that FairDolce substantially outperforms alternative approaches with robust performance under dynamic environments in achieving the highest accuracy of all time.

As a tough competitor, FairSAOML addresses the same problem that we stated in this paper by proposing an expert-tracking technique in which experts' weights are updated accordingly. It assumes that larger experts containing information across a large number of tasks help the learner to adapt to the new environment quickly. In our experiments, although FairSAOML shows competitive performance in bias control, it cannot surpass ours. This is because when the environment changes, larger experts in FairSAOML retain information from the old environments, which hurts the performance of the learner. Similar reasons are attributed to interval-based learning algorithms, such as AOD and CBCE. In the case of changing environments, one major merit of FairDolce is to disentangle data by separated representations in latent spaces, where only the semantic ones correspond to model predictions. This effectively controls the interference from various environments.

7.2 Ablation Studies

We conduct ablation studies on all datasets to demonstrate the contributions of three key components in our method. Fig. 4 demonstrates the results on the rcMNIST dataset. Results on other datasets refer to Figs. 5 to 7 in Appendix B. (1) In this first study ($w/o\, h_v \& D$), we intentionally remove the variation encoder h_v and the decoder D and only keep the semantic encoder h_s and the classifier ω with fair constraints. In this sense, the proposed architecture is equivalent to a simple neural network, and the semantic encoder functions as a featurizer. (2) In the second study (w/o fair constraints), we keep all modules but remove the fairness constraints g from the classifier ω. Without fair constraints, although the model provides better performance, fairness is not guaranteed over time. (3) In the third study ($w/o\, h_v$), only the variation encoder h_v is removed. Without the variation encoder, the model is similar to conventional autoencoders. The generalization ability to changing environments is weakened.

8 CONCLUSION

To address the problem of fairness-aware online learning for changing environments, we first introduce a novel regret, namely FairSDR, in which it takes a mixed form of static and dynamic regret metrics. We challenge existing online learning methods by sequentially updating model parameters with a local change, where only parts of

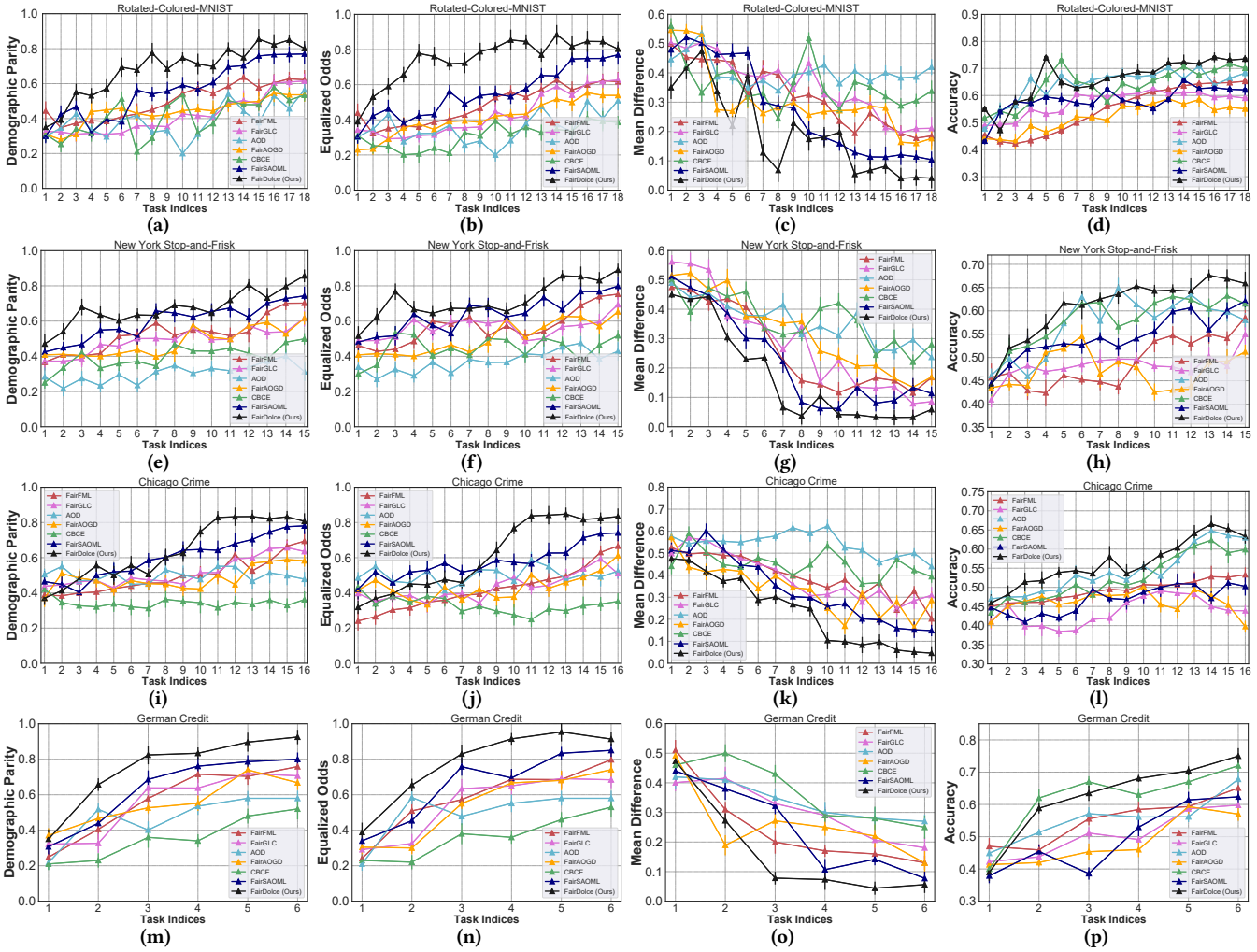

Figure 3: Model performance over datasets through each time. (a-d) Rotated-Colored-MNIST; (e-h) New York Stop-and-Frisk, (i-l) Chicago Crime; (m-p) German Credit.

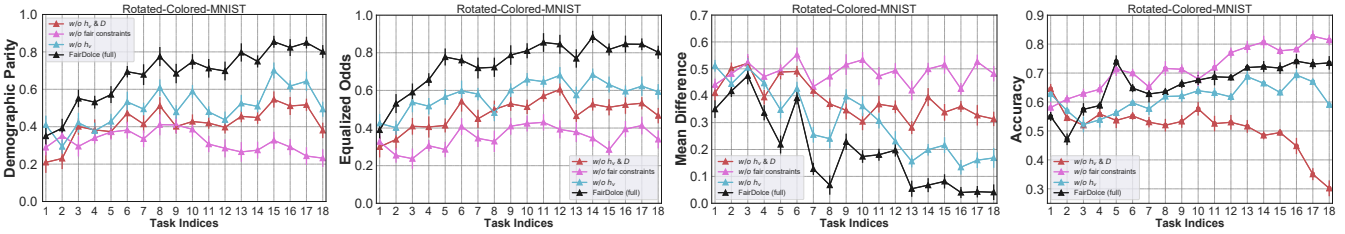

Figure 4: Ablation studies on the Rotated-Colored-MNIST dataset.

the parameters correspond to environmental change, and keep the remaining invariant to environments and thus solely for fair predictions. To this end, an effective algorithm FairDolce is introduced, wherein it consists of two networks with auto-encoders. Through disentanglement, data are able to be encoded with an environment-invariant semantic factor and an environment-specific variation factor. Furthermore, semantic factors are used to learn a classifier under a group fairness constraint. Detailed theoretic analysis and corresponding proofs justify the effectiveness of the proposed algorithm by demonstrating upper bounds for the loss regret and violation of fair constraints. Empirical studies based on real-world datasets show that our method outperforms state-of-the-art online learning techniques in both model accuracy and fairness.

ACKNOWLEDGMENTS

The research reported was supported by the National Science Foundation under grant number 2147375 and 1750911.

REFERENCES

[1] Arthur Asuncion and David Newman. 2007. In *UCI machine learning repository*.

[2] Dan Biddle. 2005. Adverse Impact and Test Validation: A Practitioner's Guide to Valid and Defensible Employment Testing. *Gower* (2005).

[3] Amit Daniely, Alon Gonen, and Shai Shalev-Shwartz. 2015. Strongly Adaptive Online Learning. In *ICML*.

[4] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Rich Zemel. 2011. Fairness Through Awareness. *CoRR* (2011).

[5] Michael Feldman, Sorelle Friedler, John Moeller, Carlos Scheidegger, and Suresh Venkatasubramanian. 2015. Certifying and Removing Disparate Impact. *KDD* (2015).

[6] Chelsea Finn, Aravind Rajeswaran, Sham Kakade, and Sergey Levine. 2019. Online Meta-Learning. *ICML* (2019).

[7] Muhammad Ghifary, W Bastiaan Kleijn, Mengjie Zhang, and David Balduzzi. 2015. Domain generalization for object recognition with multi-task autoencoders. In *Proceedings of the IEEE international conference on computer vision*. 2551–2559.

[8] András Gyorgy, Tamás Linder, and Gábor Lugosi. 2012. Efficient tracking of large classes of experts. *IEEE Transactions on Information Theory* (2012).

[9] James Hannan. 1957. Approximation to bayes risk in repeated play. *Contributions to the Theory of Games* (1957).

[10] Moritz Hardt, Eric Price, and Nathan Srebro. 2016. Equality of opportunity in supervised learning. *NeurIPS* (2016).

[11] Elad Hazan and C. Seshadhri. 2007. Adaptive Algorithms for Online Decision Problems. *Electronic Colloquium on Computational Complexity (ECCC)* (2007).

[12] Sergey Ioffe and Christian Szegedy. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*. pmlr, 448–456.

[13] Rodolphe Jenatton, Jim Huang, and Cedric Archambeau. 2016. Adaptive Algorithms for Online Convex Optimization with Long-term Constraints. *ICML* (2016).

[14] Kwang-Sung Jun, Francesco Orabona, Stephen Wright, and Rebecca Willett. 2017. Improved Strongly Adaptive Online Learning using Coin Betting. In *AISTATS*.

[15] Pang Wei Koh, Shiori Sagawa, Henrik Marklund, Sang Michael Xie, Marvin Zhang, Akshay Balsubramani, Weihua Hu, Michihiro Yasunaga, Richard Lanas Phillips, Irena Gao, Tony Lee, Etienne David, Ian Stavness, Wei Guo, Berton Earnshaw, Imran Haque, Sara M Beery, Jure Leskovec, Anshul Kundaje, Emma Pierson, Sergey Levine, Chelsea Finn, and Percy Liang. 2021. WILDS: A Benchmark of in-the-Wild Distribution Shifts. In *ICML*.

[16] Ming-Yu Liu, Thomas Breuel, and Jan Kautz. 2017. Unsupervised image-to-image translation networks. *Advances in neural information processing systems* 30 (2017).

[17] Michael Lohaus, Michael Perrot, and Ulrike Von Luxburg. 2020. Too Relaxed to Be Fair. In *ICML*.

[18] Mehrdad Mahdavi, Rong Jin, and Tianbao Yang. 2012. Trading regret for efficiency: online convex optimization with long term constraints. *JMLR* (2012).

[19] Changdae Oh, Heeji Won, Junhyuk So, Taero Kim, Yewon Kim, Hosik Choi, and Kyungwoo Song. 2022. Learning Fair Representation via Distributional Contrastive Disentanglement. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1295–1305.

[20] Vishakha Patil, Ganesh Ghalme, Vineet Nair, and Yadati Narahari. 2020. Achieving Fairness in the Stochastic Multi-Armed Bandit Problem. *AAAI* (2020).

[21] Alexander Robey, George J Pappas, and Hamed Hassani. 2021. Model-based domain generalization. *Advances in Neural Information Processing Systems* 34 (2021), 20210–20229.

[22] Bernhard Schölkopf, Francesco Locatello, Stefan Bauer, Nan Rosemary Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. 2021. Toward causal representation learning. *Proc. IEEE* 109, 5 (2021), 612–634.

[23] Yuanyu Wan, Bo Xue, and Lijun Zhang. 2021. Projection-free Online Learning in Dynamic Environments. *AAAI* (2021).

[24] Yongkai Wu, Lu Zhang, and Xintao Wu. 2019. On Convexity and Bounds of Fairness-aware Classification. *WWW*.

[25] Tianbao Yang, Lijun Zhang, Rong Jin, and Jinfeng Yi. 2016. Tracking slowly moving clairvoyant: Optimal dynamic regret of online learning with true and noisy gradient. In *ICML*.

[26] Xinlei Yi, Xiuxian Li, Tao Yang, Lihua Xie, Tianyou Chai, and Karl Johansson. 2021. Regret and cumulative constraint violation analysis for online convex optimization with long term constraints. In *International Conference on Machine Learning*. PMLR, 11998–12008.

[27] Hao Yu, Michael J. Neely, and Xiaohan Wei. 2017. Online Convex Optimization with Stochastic Constraints. In *NeurIPS*.

[28] Jianjun Yuan and Andrew Lamperski. 2018. Online convex optimization for cumulative constraints. *NeurIPS* (2018).

[29] Richard Zemel, Yu Wu, Kevin Swersky, Toniann Pitassi, and Cynthia Dwork. 2013. Learning Fair Representations. *ICML* (2013).

[30] Hanlin Zhang, Yi-Fan Zhang, Weiyang Liu, Adrian Weller, Bernhard Schölkopf, and Eric P Xing. 2022. Towards principled disentanglement for domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8024–8034.

[31] Lijun Zhang, Shiyin Lu, and Tianbao Yang. 2020. Minimizing Dynamic Regret and Adaptive Regret Simultaneously. *AISTATS* (2020).

[32] Lijun Zhang, Shiyin Lu, and Zhi-Hua Zhou. 2018. Adaptive Online Learning in Dynamic Environments, In International Conference on Neural Information Processing Systems. *NeurIPS 2018*.

[33] Lijun Zhang, Tianbao Yangt, Jinfeng Yi, Rong Jin, and Zhi-Hua Zhou. 2017. Improved Dynamic Regret for Non-Degenerate Functions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Red Hook, NY, USA, 732–741.

[34] Chen Zhao and Feng Chen. 2019. Rank-Based Multi-task Learning For Fair Regression. *IEEE International Conference on Data Mining (ICDM)* (2019).

[35] Chen Zhao and Feng Chen. 2020. Unfairness Discovery and Prevention For Few-Shot Regression. *ICKG* (2020).

[36] Chen Zhao, Feng Chen, and Bhavani Thuraisingham. 2021. Fairness-Aware Online Meta-learning. *ACM SIGKDD* (2021).

[37] Chen Zhao, Feng Mi, Xintao Wu, Kai Jiang, Latifur Khan, and Feng Chen. 2022. Adaptive Fairness-Aware Online Meta-Learning for Changing Environments. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2565–2575. https://doi.org/10.1145/3534678.3539420

[38] Martin Zinkevich. 2003. Online Convex Programming and Generalized Infinitesimal Gradient Ascent. *ICML* (2003).

A ADDITIONAL EXPERIMENT DETAILS

A.1 Notations

Vectors are denoted by lowercase bold face letters. Scalars are denoted by lowercase italic letters. Sets are denoted by uppercase calligraphic letters. Indices of task sequences are denoted as $[T] = \{1, \cdots, T\}$. $||\cdot||$ represents ℓ_2 norm.

Table 2: Important notations and corresponding descriptions.

Notations	Descriptions
T	total number of learning tasks
t	indices of tasks
f_t	loss function at time t
g	fairness function
ω	classification function
h_s	semantic encoder
h_v	variation encoder
D	decoder
$\boldsymbol{\theta}$	model parameters
$\boldsymbol{\theta}^s$	parameters of the semantic encoder
$\boldsymbol{\theta}^v$	parameters of the variation encoder
$\boldsymbol{\theta}^d$	parameters of the decoder
$\boldsymbol{\theta}^{cls}$	parameters of the classifier
$\mathbf{u}^s$	semantic comparators
$\mathbf{s}$	semantic factor (representation)
$\mathbf{v}$	variation factor
$\mathcal{Q}_t$	data batch sampled from the task pool at time t
Q	total number of quartet/doublet pairs in $\mathcal{Q}_t$
q	indices of quartet/doublet pair in $\mathcal{Q}_t$
$\lvert\mathcal{Q}_t\rvert$	total number of samples in the batch $\mathcal{Q}_t$

A.2 Hyperparameter Search

For each dataset, we tune the following hyperparameters: (1) the initial dual meta parameter $\lambda_1, \lambda_2, \lambda_3$ is chosen from {0.00001, 0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000, 10000 }; (2) learning rates η_1 and η_2 for updating primal and dual variables are chosen from {0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 50, 100, 500, 1000}; (3) set margins $\eta_1 = \eta_2 = \eta_3 = 0.05$.

B ADDITIONAL EXPERIMENT RESULTS

Ablation study results on the New York Stop-and-Frisk, Chicago Crime, and German Credit datasets are shown in Figs. 5 to 7. Similar trends are observed as the Rotated-Colored-MNIST in Fig. 4.

C SKETCH PROOF OF THEOREM 1

Proof. Using $\eta_{1,t} = \frac{\eta_{1,0}}{\sqrt{T}}, \eta_{2,t} = \frac{\eta_{2,0}}{\sqrt{\eta_{1,t}}}, \forall t \in [T]$ given in the Theorem yelds

$$\sum_{t=1}^{T} \Delta_t(\mathbf{u}_T^c) = \frac{\sqrt{T}}{\eta_{1,0}} \sum_{t=1}^{T} (||\mathbf{u}_T^c - \boldsymbol{\theta}_t^l||^2 - ||\mathbf{u}_T^c - \boldsymbol{\theta}_{t+1}^l||^2)$$
$$\leq \frac{\sqrt{T}}{\eta_{1,0}} ||\mathbf{u}_T^c - \boldsymbol{\theta}_1^l||^2$$

Combing the above inequality with the Lemma 1 presented in [26] yields

$$\sum_{t=1}^{T} f_t(h_s(\boldsymbol{\theta}_t^s), \boldsymbol{\theta}_t^{cls}) - \min_{\boldsymbol{\theta}^{cls} \in \Theta} \sum_{t=1}^{T} f_t(h_s(\mathbf{u}_t^s), \boldsymbol{\theta}^{cls})$$
$$\leq \frac{\sqrt{T}}{\eta_{1,0}} ||\mathbf{u}_T^c - \boldsymbol{\theta}_1^l||^2 + \frac{G^2 \eta_{1,0}}{2} \sqrt{T}$$

which yields the bound for the loss regret. Similarly, with the Lemma 1 presented in [26], it yields

$$||\lambda_{T,1}||^2 \leq \beta T$$

where $\beta = \frac{2}{\eta_{1,0}\sqrt{T}} ||\mathbf{u}_T^c - \boldsymbol{\theta}_1^l||^2 + \frac{G^2 \eta_{1,0}}{\sqrt{T}} + \frac{\eta_{2,0}^2 F^2}{\eta_{1,0} T T} + 2F$. Together the above inequality with $\eta_{1,t} = \frac{\eta_{1,0}}{\sqrt{T}}, \eta_{2,t} = \frac{\eta_{2,0}}{\sqrt{\eta_{1,t}}}, \forall t \in [T]$, we have

$$\sum_{t=1}^{T} \left\| \left[g(h_s(\boldsymbol{\theta}_t^s), \boldsymbol{\theta}_t^{cls}) \right]_+ \right\| \leq \frac{\sqrt{m \eta_{1,0}}}{\eta_{2,0} T} ||\lambda_{T,1}|| \leq \frac{m \eta_{1,0} \beta}{\eta_{2,0}} T^{\frac{1}{4}}$$

which yields the bounds for the violation of the long-term constraints. □

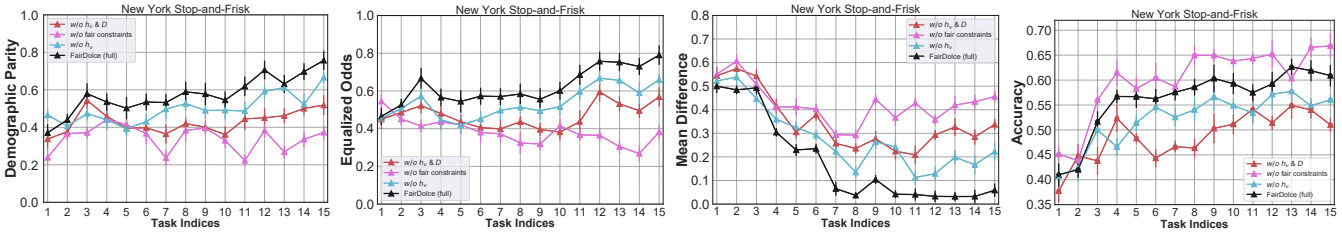

Figure 5: Ablation studies on the New York Stop-and-Frisk dataset.

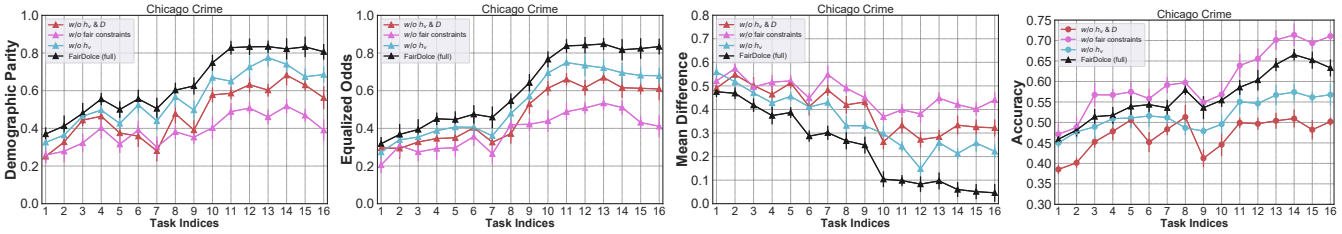

Figure 6: Ablation studies on the Chicago Crime dataset.

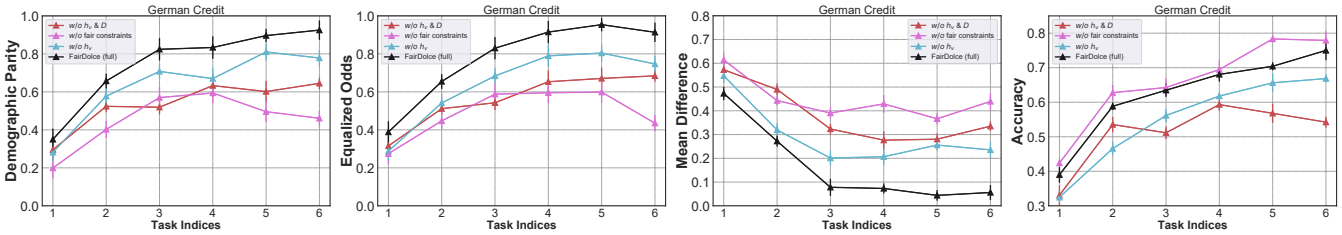

Figure 7: Ablation studies on the German Credit dataset.