

Fairness In a Non-Stationary Environment From an Optimal Control Perspective

Anonymous Authors¹

Abstract

The performance of state-of-the-art machine learning models is observed to degrade in scenarios involving under-represented demographic populations during training. This issue has been extensively studied within a supervised learning framework where data distribution remains unchanged. Nonetheless, real-world use cases often encounter distribution shifts induced by the models in deployment. For example, performance bias against minority users can affect customer retention rates, thereby skewing available data from active users due to the absence of minority user input. This feedback effect further exacerbates the discrepancy across various demographic groups in subsequent time steps. To mitigate this problem, we introduce asymptotic fairness, a criterion that aims at preserving sustained model performance across all demographic populations. In addition, we construct a surrogate retention system, based on existing literature on evolutionary population dynamics, to approximate the dynamics of distribution shifts on active user counts. This system allows the aim of achieving asymptotic fairness to be formulated as an optimal control problem. To evaluate the effectiveness of the proposed method, we design a generic simulation environment that simulates the population dynamics of the feedback effect between user retention and model performance. When we deploy the models to this simulation environment, by considering long-term planning, the optimal control solution outperforms existing baseline methods, demonstrating superior performance.

¹Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

1. Introduction

In dynamically changing environments, data distributions can evolve over time, giving rise to the phenomenon known as concept drift. Concept drift entails changes in the conditional distribution of the target variable given the input features, while the distribution of the input features themselves might remain fixed (Gama et al., 2014; Schlimmer & Granger, 1986; Widmer & Kubat, 1996). This work considers a distinct case of concept drift, where the performance of a machine learning system can impact the number of active users, while the feature distribution of these users remains fixed. A relevant scenario involves users interacting with personal devices such as Google Home or Amazon Alexa, providing input features like voice commands. Device performance feedback, be it positive (accurate voice identification) or negative (misidentification), can respectively increase user engagement or decrease user retention. This issue becomes more pronounced when populations from different demographic groups are considered (Harwell). A system demonstrating performance bias against minority demographics can result in diminished retention rates for these users. Subsequently, the available training data collected from active users might be insufficient to represent the distribution of minority users. Consequently, this can further amplify the representation disparity issue when the model is fine-tuned based on this training set (Hashimoto et al., 2018).

In this study, we aim to develop machine learning models that engage users from all demographic groups, under the aforementioned population dynamics. This task can be viewed as a trajectory planning problem, wherein the evolution of user engagement is optimized through control design. With the given population dynamics, trajectory planning can be considered as an optimal control problem, solvable by existing methods, such as dynamic programming (Bellman, 1952). However, the generation and evaluation of an optimal control solution pose three challenges. Firstly, the notion of fairness in a non-stationary environment has not been formally defined. Existing fairness definitions aim to measure performance disparities between different demographic groups at a singular time point, such as equal opportunity (Hardt et al., 2016), and demographic parity (Feldman et al., 2015). These definitions fail to adequately encapsulate the goals in a non-stationary environment. Second, solving the

optimal control problem requires knowledge of the underlying population dynamics, which are typically complex and inaccessible. Lastly, evaluation of the performance of an optimal control solution (e.g., one generated from some estimated dynamic system) necessitates its deployment on population dynamics with real-world users, an often costly and unattainable requirement. In response to these challenges, our contributions are threefold:

1. We introduce the concept of asymptotic fairness to define the consistent performance across all demographic groups over an extended time period.
2. We propose a surrogate retention system built upon the existing literature on evolutionary population dynamics, from which we formulate the objective of achieving asymptotic fairness as an optimal control problem. To address this control problem, we consider Pontryagin’s maximum principle which allows us to solve for the optimal control solution efficiently.
3. Through empirical evaluation, we highlight the advantages of taking into account the underlying dynamics in model design. Our results consistently outperform existing baseline methods, thereby validating the superiority of our approach in terms of performance.

Furthermore, we design a simulator that simulates the non-stationary environment of the user’s willingness to retain or churn from a deployed model. This simulator allows for testing the evolutionary fairness property of machine learning models in synthetic population dynamics.

2. Related Works

We review existing literature about fairness in non-stationary settings, and further discuss the intersection of machine learning and optimal control, emphasizing its relevance and applicability to our work.

Fairness problems in the non-stationary setting. Recent studies have brought attention to the potential pitfalls of imposing static fairness constraints (Hardt et al., 2016; Feldman et al., 2015). These constraints can yield unfavorable long-term effects, as demonstrated in the work by (Liu et al., 2018; Zhang et al., 2020). The interplay between algorithmic decisions and individuals’ reactions plays a pivotal role in shaping these long-term effects (Zhang et al., 2020). For instance, model decisions can cause changes in the underlying data distribution, subsequently affecting the model’s performance in future time steps. (Zhang et al., 2019) recently presented a comprehensive study of the interaction between user retention rates and model decisions in dynamic environments. The most common approach to address this problem is via successive one-step methods,

which prioritize fairness for minority demographic groups (Hashimoto et al., 2018). In view of these insights, our research emphasizes the importance of considering the inherent dynamics within a non-stationary environment.

The connection between deep learning and optimal control. Recent works have highlighted the connection between dynamical systems and deep neural networks (E, 2017; Haber & Ruthotto, 2017). This perspective offers valuable theoretical insights for understanding deep learning through an optimal control lens (Liu & Theodorou, 2019). The pioneering work that bridged and extended the classical back-propagation algorithm with optimal control theory was introduced by (Li et al., 2018; Li & Hao, 2018), establishing a direct relationship between the Pontryagin’s maximum principle (Kirk, 1970) and gradient-based training. Building upon this foundation, (E et al.) established the mathematical basis for the optimal control viewpoint in deep learning. Moreover, optimal control methods have been applied to tackle challenging issues in deep learning. For instance, (Liu et al., 2020a;b) proposed efficient high-order optimizers using differential dynamic programming, and (Chen et al., 2021) explored closed-loop controllers to improve robustness against adversarial attacks. These advances highlight the efficacy of optimal control approaches in resolving key challenges of deep learning.

3. Fairness in Non-Stationary Environment

In this section, we discuss the problem setup for fairness in a non-stationary environment, where user retention or churn is conditioned on the model’s performance on the data they provide. This configuration leads to a condition for the machine learning models, which we term as asymptotic fairness. This condition requires the models to sustain their performance across all demographic groups over an extended period.

3.1. Problem Description for Fairness in a Non-Stationary Environment

In a non-stationary environment, we consider a predictive model with time-varying model parameters, denoted as $\{\theta_t\}_{t=0}^{T-1}$. We consider K distinct demographic groups, each comprising N^i users that include both participative and non-participative users of the predictive model. Let us denote Λ_t^i as the number of active users within the i^{th} demographic group at time t . To facilitate our analysis, we normalize Λ_t^i by defining a population density $\lambda_t^i = \frac{\Lambda_t^i}{N^i}$, which falls within the range of 0 to 1. The growth (resp. decay) of λ_t^i indicates that more users from the i^{th} demographic group participate in (resp. leave) the system. We denote $\mathcal{A}_t^i$ as a N -tuple to indicate the active users from the i^{th} demographic group at time step t (e.g. $\mathcal{A}_t^i = (1, 0, 1)$)

means that the 1st and 3rd users from the i^{th} group are active at time step t). We formulate the population dynamics of the user's willingness to participate as a Markov decision process and denote it as **population retention system**,

$$\mathcal{A}_{t+1}^i \sim \mathcal{P}(\cdot | \mathcal{A}_t^i, \theta_t, \{\mathbf{z}^{i,n}\}_{n=1}^N, \Phi(\cdot)), \quad (1)$$

where the state space is a N -tuple that indicates the indices of active users, $\mathcal{P}(\cdot)$ denotes the transition probability, θ_t is the model parameter at time step t , $\mathbf{z}^{i,n}$ represents the features of the n^{th} user from the i^{th} demographic group, this feature vector contains both input data (e.g. voice command) and characteristics (e.g. the probability of a user churn when encountered with a correct prediction, and the probability of a user retaining when receiving a wrong prediction) of the user, $\Phi(\cdot)$ is a loss (or reward) function acting on each individual user. In addition, a certain number of inactive users become active in the system at each time step, this number is conditioned on the model performance of currently active users. In a nutshell, if a user is active at time step t , the activeness of this user at the next time step is conditioned on the model performance of that specific user. Moreover, data features of only active users at each time step are considered an observable state that can be leveraged for generating a model. Once a sequence of models $\{\theta_t\}_{t=1}^T$ is generated, we evaluate its performance using the population retention system defined in Eq. (1).

3.2. The Definition for Asymptotic Fairness

The population retention system, defined in Eq. (1), simulates variation in the active user population density, resulting from the individual model performance of each user. In this context, the deployment of a model with an adequately minimized population risk encourages increased user participation within the system. This encourages more active users to subsequently contribute training data for model refinement, culminating in an enhanced model, which in turn can further decrease population risk in the subsequent time step. This positive feedback loop, given enough time, results in the population risks for all demographic groups converging towards 0 for a large number of total users N^i , while the number densities approach 1. This motivates us to define the concept of asymptotic fairness as follows:

Definition 3.1. Asymptotic fairness A sequence of models satisfies asymptotic fairness if the dynamics it drives satisfy the following condition:

$$\begin{aligned} \lambda_t^i &\rightarrow 1, \text{ as } t \rightarrow \infty \quad \forall i \in [1, 2, \dots, K], \\ \text{s.t. } \mathcal{A}_{t+1}^i &\sim \mathcal{P}(\cdot | \mathcal{A}_t^i, \theta_t, \{\mathbf{z}^{i,n}\}_{n=1}^N, \Phi(\cdot)), \\ \lambda_t^i &= \frac{\sum_{n=1}^{N^i} [\mathcal{A}_t^i]_n}{N^i}. \end{aligned}$$

Furthermore, the satisfaction of asymptotic fairness by a sequence of models is implicitly linked to the initial popula-

tion densities. In scenarios where all demographic groups initially have high population densities, the likelihood of achieving asymptotic fairness increases. Conversely, scenarios with highly imbalanced representations of demographic groups pose significant challenges in meeting this condition. Therefore, the representation of demographic groups plays a critical role in the successful implementation of models adhering to the condition of asymptotic fairness.

Remark 3.2. The mode of convergence as per Definition 3.1 is application-dependent, thus, it varies across different contexts. In this work, we introduce a deterministic surrogate system as defined in Section 4.1. Given the deterministic nature of this surrogate system, we abstain from specifying the mode of convergence pertaining to random variables. The absence of inherent stochasticity in the surrogate system obviates the need for a specific convergence criterion that would otherwise be necessary for the presence of random variables.

Remark 3.3. The concept of asymptotic fairness provides a more precise interpretation of fairness in a non-stationary environment. Prior research has considered disparity amplification (Hashimoto et al., 2018) to assess the representation disparity across all demographic groups at each individual time step. However, the definition of asymptotic fairness diverges from this approach as it emphasizes long-term behavior. To illustrate, consider an extreme scenario where the population densities of all demographic groups concurrently decay to zero. Although this is an undesirable situation, it would nonetheless satisfy the condition of disparity amplification, yet not meet the criterion of asymptotic fairness. Thus, the distinction underscores the importance of considering long-term behavior in fairness definitions, a perspective that asymptotic fairness uniquely encapsulates.

4. An Optimal Control Solution for Asymptotic Fairness

According to Def. 3.1, our goal is to maximize the population densities across all demographic groups within the context of the population retention system, as defined by Eq. (1). Due to the inaccessibility of the underlying dynamics of the population retention system, our initial step involves the construction of a surrogate system to estimate these dynamics. Subsequently, we formulate the condition of asymptotic fairness as an optimal control problem and provide an efficient solver based on Pontryagin's maximum principle (PMP) (Pontryagin, 1987).

4.1. Surrogate Retention System for the Evolutionary Population Dynamics

Our design of the surrogate retention system is rooted in the existing body of literature on evolutionary population dynamics (Cushing, 2019). This system features a low-

dimensional state representation, adding to its computational efficiency. Moreover, the system is deterministic and differentiable, which allows for solving the optimal control solution efficiently. Thus, the surrogate retention system not only provides a meaningful connection to evolutionary dynamics but also offers practical advantages in terms of computational efficiency.

Evolutionary population dynamics describes the dynamics of user participation. Difference equations are commonly employed to describe discrete-time dynamics where temporal variations in vital rates are influenced by population density dependencies. Activities of an individual, such as reproduction and survival, may experience fluctuations, thereby contributing to the evolutionary dynamics of population density. Such explicit temporal dependencies can be encapsulated through the optimization of the coefficients of a difference equation over time (Vincent & Brown, 2005). To account for such evolutionary mechanisms, a difference equation population model can be developed (Cushing, 2019). In a simplified context, the population's growth and decay are attributed to births and deaths, respectively. Individuals present at time step $t + 1$ are either newcomers within the time frame or survivors from time step t . We model those dynamics as the following discrete dynamic system:

$$\lambda_{t+1}^i = (1 - \lambda_t^i)\beta(\kappa^i(\lambda_t^i, \theta_t)) + \lambda_t^i\sigma(\kappa^i(\lambda_t^i, \theta_t)), \quad (2)$$

where $\kappa^i(\cdot)$ computes a value that reflects the response of the i^{th} population on some external controls θ_t (e.g. medical treatment, resource allocation), $\beta(\cdot)$ and $\sigma(\cdot)$ compute the ratios of newborns and survived population during a time interval respectively. We consider the range of both birth and survival rate functions to be $[0, 1]$, in which case, the population densities take the range of $[0, 1]$. Moreover, we denote $\lambda_t = [\lambda_t^1, \lambda_t^2, \dots, \lambda_t^K]^T$ as a K -dimensional vector. Subsequently, a K -dimensional discrete dynamic system describing the simplified evolutionary population dynamics can be constructed as follows:

$$\lambda_{t+1} = \mathcal{T}(\lambda_t, \theta_t), \quad (3)$$

where the i^{th} element of $\mathcal{T}(\cdot)$ is defined in Eq. (2).

In cases where user retention or churn rates are influenced by population dynamics, the function $\kappa^i(\cdot)$ is used to measure the model's performance on the currently active population. Here, the birth rate $\beta(\cdot)$ and the survival rate $\sigma(\cdot)$ denote the proportions of incoming and retained users at each respective time step. For performance evaluation of the model, we maintain a small holdout set of accessible users, all of whose user features are observable. From this set, data from active users can be sampled at every time step, based on the population density λ_t^i . Moreover, when model performance is evaluated via a reward (or alternately, a loss) function, we

hypothesize that the birth and survival rates correspond proportionally (or inversely) to the model performance $\kappa^i(\cdot)$. This assumption ensures that an improved model performance leads to an increase in population density. We denote this difference equation as the **surrogate retention system**.

Model evaluation through the worst-case distributional loss. The surrogate retention system, as defined by Eq. (3), leads to a low-dimensional state representation, consisting only of the population densities across all demographic groups. However, its simulation requires random sampling of training data from the holdout set, based on the present population density. This is inconsistent with the population retention system, where the training data are provided from active users. To resolve this discrepancy, we consider the formulation of distributionally robust optimization, which considers the λ_t^i proportion of users who received optimal model performance.

To begin with, let $d_{\mathcal{X}^2}(\mathcal{M}||\mathcal{Q}) = \int (\frac{d\mathcal{M}}{d\mathcal{Q}} - 1)^2 d\mathcal{Q}$ denote the $\mathcal{X}^2$ -divergence between two probability distributions $\mathcal{M}$ and $\mathcal{Q}$, $\mathcal{B}(\mathcal{M}, r) = \{\mathcal{Q} : d_{\mathcal{X}^2}(\mathcal{M}||\mathcal{Q}) \leq r\}$ denote the chi-squared ball around a probability distribution $\mathcal{M}$ of radius r . Let $\mathcal{M}^i$ be the feature distribution of users from the i^{th} demographic group, we consider the performance measure $\kappa^i(\cdot)$ as the worst-case distributional loss over all r -radius balls around $\mathcal{M}^i$ defined as follows,

$$\begin{aligned} \kappa^i(\lambda_t^i, \theta_t) &= \sup_{\mathcal{Q} \in \mathcal{B}(\mathcal{M}^i, r_t^i)} \mathbb{E}_{\mathbf{z} \sim \mathcal{Q}} \Phi(\theta_t, \mathbf{z}), \\ r_t^i &= (1/\lambda_t^i - 1)^2. \end{aligned} \quad (4)$$

Clearly, as the number density λ_t^i approaches 1, r_t^i decays to 0, and $\kappa^i(\lambda_t^i, \theta_t)$ is equivalent to population risk. For small λ_t^i , the radius $r_t^i \rightarrow \infty$ and this leads to a large loss value. In general, computing the worst-case distributional loss over a set of distributions is a challenging task. Fortunately, the maximization problem in Eq. (4) can be reformulated into its dual form (Duchi et al., 2019). More specifically, if $\Phi(\cdot)$ is upper semi-continuous for any θ , then for $r_t^i \geq 0$ and any θ , the following holds true:

$$\begin{aligned} &\sup_{\mathcal{Q} \in \mathcal{B}(\mathcal{M}^i, r_t^i)} \mathbb{E}_{\mathbf{z} \sim \mathcal{Q}} \Phi(\theta_t, \mathbf{z}) \\ &= \inf_{\eta \in \mathbb{R}} \left(C(\lambda_t^i) \cdot (\mathbb{E}_{\mathcal{M}^i} [\Phi(\theta_t, \mathbf{z}) - \eta]_+^2)^\frac{1}{2} + \eta \right), \\ &\text{where } C(\lambda_t^i) = (2(1/\lambda_t^i - 1)^2 + 1)^\frac{1}{2}, \end{aligned} \quad (5)$$

where $[x]_+ = x$ if $x \geq 0$ and 0 otherwise. At time step t , given θ_t and λ_t^i , the worst-case distributional loss is computed by averaging the sample losses that are higher than the optimal $\eta^*(\lambda_t^i, \theta_t)$, where $\eta^*(\lambda_t^i, \theta_t)$ attains the infimum.

Rather than calculating the worst-case distributional loss, we consider $\Phi(\cdot)$ as a reward function. In this setting, the worst-case distributional loss corresponds to the computation of

the maximal distributional reward, as defined within the chi-square ball of radius λ_t^i , and the active users are the ones who received a reward equal to or greater than $\eta^*(\lambda_t^i, \theta_t)$. This computation can be performed efficiently using Eq. (5).

Overall, the surrogate retention system for the i^{th} demographic population is defined as follows,

$$\begin{aligned} \lambda_{t+1}^i &= (1 - \lambda_t^i)\beta(\kappa^i(\lambda_t^i, \theta_t)) + \lambda_t^i\sigma(\kappa^i(\lambda_t^i, \theta_t)), \\ \text{where } \kappa^i(\lambda_t^i, \theta_t) &= \sup_{\mathcal{Q} \in \mathcal{B}(\mathcal{M}^i, r_t^i)} \mathbb{E}_{\mathbf{z} \sim \mathcal{Q}} \Phi(\theta_t, \mathbf{z}), \\ r_t^i &= (1/\lambda_t^i - 1)^2. \end{aligned}$$

Here we describe two remarks about the proposed surrogate retention system.

Remark 4.1. The surrogate retention system, when implemented with the worst-case distributional loss formulation, yields a low-dimensional state representation, with the state consisting of the population densities of all demographics. Furthermore, given a model θ_t and population density λ_t , the system represents a deterministic difference equation where active users are generated as per Eq.(5). The system is differentiable with respect to the state λ_t , allowing for efficient computation of the optimal model parameters, as detailed in Sec.4.2.

Remark 4.2. Model-based reinforcement learning (RL) techniques offer notable advantages in terms of sample complexity, addressing a key challenge faced by model-free RL methods such as policy gradient-based approaches. Despite these benefits, scaling model-based RL to high-dimensional problem domains remains a challenge. In this study, we propose a novel approach that considers population density as a statistical average, effectively capturing underlying evolutions. By adopting this perspective, we simplify the problem to a low-dimensional surrogate retention system. The simulations conducted in this work can serve as a valuable foundation for the development of model-free RL methods.

4.2. Optimal Control Formulation for Asymptotic Fairness

We denote $\Psi(\cdot, \mathbf{1})$ as the binary cross-entropy loss between population densities and a fixed vector of $\mathbf{1}$ at a specific time step and $\mathbf{1}$. In this case, minimizing $\Psi(\cdot, \mathbf{1})$ is equivalent to maximizing the population densities at the terminal time step. Consequently, the objective of realizing asymptotic fairness can be formulated as follows:

$$\min_{\{\theta_t\}_{t=0}^{T-1}} \Psi(\lambda_T, \mathbf{1}) \text{ s.t. } \lambda_{t+1} = \mathcal{T}(\lambda_t, \theta_t), \text{ given } \lambda_0, \quad (6)$$

where $\mathcal{T}(\cdot)$ is the surrogate retention system defined in Eq. (3). This is a special case of a class of general optimal control problems for discrete dynamical systems, in which we consider the control variables as the model parameters at all time steps. From this optimal control perspective,

asymptotic fairness can be achieved by solving for a set of controls such that Eq. (6) is satisfied.

This is known as closed-loop control in which the optimal control $\theta_t^*(\lambda_t)$ has explicit dependence on the state. The optimal control solution can be solved via dynamical programming principle (Bellman, 1952), which involves solving a Hamilton Jacobi Bellman partial differential equation. For an efficient model generation, PMP (Pontryagin, 1987) converts dynamical programming into two difference equations and a maximization condition. Instead of computing the closed-loop control $\theta_t^*(\lambda_t)$, the PMP provides a necessary condition for the optimality with fixed control. To begin with, we define the Hamiltonian as

$$H(t, \lambda_t, \mathbf{p}_{t+1}, \theta_t) := \mathbf{p}_{t+1}^T \cdot \mathcal{T}(\lambda_t, \theta_t) - \mathcal{L}(\theta_t, \lambda_t),$$

where $\mathcal{L}(\theta_t, \lambda_t^i)$ is a running loss at time t . We consider all running losses as 0 since asymptotic fairness is defined at the terminal state. The PMP consists of a two-point boundary value problem,

$$\lambda_{t+1}^* = \nabla_p H(t, \lambda_t^{i,*}, \mathbf{p}_t^*, \theta_t^*), \quad \lambda_0 \text{ given}, \quad (7)$$

$$\mathbf{p}_t^* = \nabla_\lambda H(t, \lambda_t^{i,*}, \mathbf{p}_{t+1}^*, \theta_t^*), \quad \mathbf{p}_T = \frac{\partial \Psi(\lambda_T, \mathbf{1})}{\partial \lambda_T}, \quad (8)$$

plus a maximum condition of the Hamiltonian.

$$H(t, \lambda_t^{i,*}, \mathbf{p}_t^*, \theta_t^*) \geq H(t, \lambda_t^{i,*}, \mathbf{p}_t^*, \theta_t), \quad \forall \theta_t \text{ and } t. \quad (9)$$

We consider the method of successive approximation (Chen et al., 2022) to solve for the control solution. Given a initial condition λ_0 , Eq.(7) corresponds to the surrogate retention system. We then set a terminal condition, $\mathbf{p}_T = \frac{\partial \Psi(\lambda_T, \mathbf{1})}{\partial \lambda_T}$. During backpropagation of the adjoint state, the current state λ_t and adjoint state $\mathbf{p}_{t+1}$ remain fixed, allowing for the maximization of the t^{th} Hamiltonian for *InnerItr* iterations. This strategy permits multiple updates to the model θ_t within a single forward-backward propagation. Once a locally optimal model is achieved, we continue the back-propagation of the adjoint state $\mathbf{p}_{t+1}$ to $\mathbf{p}_t$ and optimize the model θ_{t-1} . In this configuration, executing the Hamiltonian dynamics n times can be decomposed into *maxItr* complete iterations and *InnerItr* local updates. The algorithm is depicted in Alg.1.

5. Numerical Experiments

In this section, we describe two simulation environments that function as the population retention system, as defined in Eq. (5.1). Subsequently, we empirically validate the proposed optimal control solution utilizing a synthetic dataset in Section 5.2, and two realistic datasets frequently employed in fairness studies, as discussed in Section 5.3.

Algorithm 1 Method of Successive Approximation.

```

input  $\lambda_0$ , learning rate  $\text{lr}$ ,  $\text{maxItr}$ ,  $\text{InnerItr}$ 
output models  $\{\theta_t\}_{t=0}^T$ 
for  $m = 1$  to  $\text{maxItr}$  do
    for  $t = 0$  to  $T-1$  do
        // Forward propagation (Eq. (7)).
         $\lambda_{t+1}^* = \nabla_p H(t, \lambda_t^{i,*}, \mathbf{p}_t^*, \theta_t^*)$ ,
    end for
    // Set terminal condition.
     $\mathbf{p}_T^* = \frac{\partial \Psi(\lambda_T, \mathbf{1})}{\partial \lambda_T}$ ,
    for  $t = T-1$  to  $0$  do
        for  $\tau = 0$  to  $\text{InnerItr}$  do
            // Compute Hamiltonian with  $\mathbf{p}_{t+1}$  and  $\lambda_t$ .
             $H(t, \lambda_t, \mathbf{p}_{t+1}, \theta_t) = \mathbf{p}_{t+1}^T \cdot \mathcal{T}(\lambda_t, \theta_t)$ ,
            // Max Hamiltonian (Eq. (9)).
             $\theta_t = \theta_t + \text{lr} \cdot \nabla_{\lambda} H(t, \lambda_t, \mathbf{p}_{t+1}, \theta_t)$ ,
        end for
        // Backward propagation (Eq. (8)).
         $\mathbf{p}_t^* = \nabla_{\lambda} H(t, \lambda_t^{i,*}, \mathbf{p}_{t+1}^*, \theta_t^*)$ ,
    end for
end for
    
```

5.1. A Generic Platform for Fairness in Non-Stationary Environment

In this work, we construct two distinct population retention systems. The first configuration is based on a Markov decision process, where the model's prediction accuracy at each time step determines user participation in the next time step. Additionally, the feature vector $\mathbf{z}^{i,n}$ describes the probabilities of user churn following a correct prediction and user retention subsequent to an incorrect prediction. This creates a more realistic simulation, as the decision to retain or churn is typically not solely dependent on the accuracy of the model's prediction. The second configuration describes user retention as conditioned on a consequence of the historical accuracy of model predictions for a specific user. For instance, a user might churn if the model has produced a set number of incorrect predictions (e.g., 3 accumulated wrong predictions) for that specific user in the past. In both configurations, we assume that the number of new users joining at each time step is positively correlated with the performance of the model. For instance, if the model demonstrates high accuracy across all active users at time t , it attracts a larger number of new users at $t+1$. This enables an increase in population density as the performance of the model improves. For each simulation run, a set of active users is randomly sampled according to a given λ_0 as initialization. In this case, simulation results in the evaluation phase are not expected to align with the training simulations, necessitating the models to exhibit robust generalization capabilities when different users are considered active during the initialization of the simulation. We denote $\mathcal{P}_1$ and $\mathcal{P}_2$

as the population retention systems in the Markovian and non-Markovian settings respectively.

For evaluation, we discuss a variety of existing baseline methods to compare against the proposed optimal control solution. We define $\mathcal{S}(\theta_t, \mathbf{z}^{i,n}) = \Phi(\theta_t, \mathbf{z}^{i,n})$ if $[\mathcal{A}_t^i]_n = 1$, and 0 otherwise.

- Empirical risk minimization (**ERM**) optimizes an average loss of all observable data,

$$\theta_t = \arg \min_{\theta} \frac{1}{\sum_{i=1}^K \sum_{n=1}^N [\mathcal{A}_t^i]_n} \sum_{i=1}^K \sum_{n=1}^N \mathcal{S}(\theta_t, \mathbf{z}^{i,n}).$$

- Minimax optimization (**Minimax**) optimizes the worst-case loss among all groups (Diana et al., 2021),

$$\theta_t = \arg \min_{\theta} \max_{i \in [1, \dots, K]} \frac{1}{\sum_{n=1}^N [\mathcal{A}_t^i]_n} \sum_{n=1}^N \mathcal{S}(\theta_t, \mathbf{z}^{i,n}).$$

- Distributionally robust optimization (**DRO**) sets a lower bound on the population density of active users and considers a proportion of data that produces the worst-case loss (Hashimoto et al., 2018).
- A greedy control (**Greedy**) is equivalent to one-step planning of the optimal control method,

$$\begin{aligned} \theta_t &= \arg \min_{\theta} \Psi(\lambda_{t+1}, \mathbf{1}) \\ \text{s.t. } \lambda_{t+1} &= \mathcal{T}(\lambda_t, \theta_t), \text{ given } \lambda_t. \end{aligned}$$

5.2. Modeling with Synthetic Dataset

In this study, we employ both population retention systems, $\mathcal{P}_1$ and $\mathcal{P}_2$, with a synthetic binary classification dataset. As depicted in Fig. 1 (a), the synthetic dataset is composed of two Gaussian blobs, each centered at disparate locations, to formulate the feature distributions of two demographic groups. The blobs located on the left and right are denoted as the majority and minority demographic groups, respectively, with respective initial population densities of 0.7 and 0.3. All experiments are reiterated with five random seeds, and the binary cross-entropy loss $\Psi(\lambda_t, \mathbf{1})$ is used to quantify the quality of each state λ_t . For instance, as λ_t increases towards $\mathbf{1}$, this loss diminishes to zero.

In the proposed optimal control methodology, we express both the survival rate and birth rate functions of the surrogate retention system as a weighted sum of polynomials. Subsequently, we impose constraints on the weighting parameters to ensure a monotonically increasing behavior in both functions, thereby aligning with the assumption underpinning the proposed surrogate retention system. During each simulation run, the optimal control method learns a surrogate retention system by optimizing the weight parameters

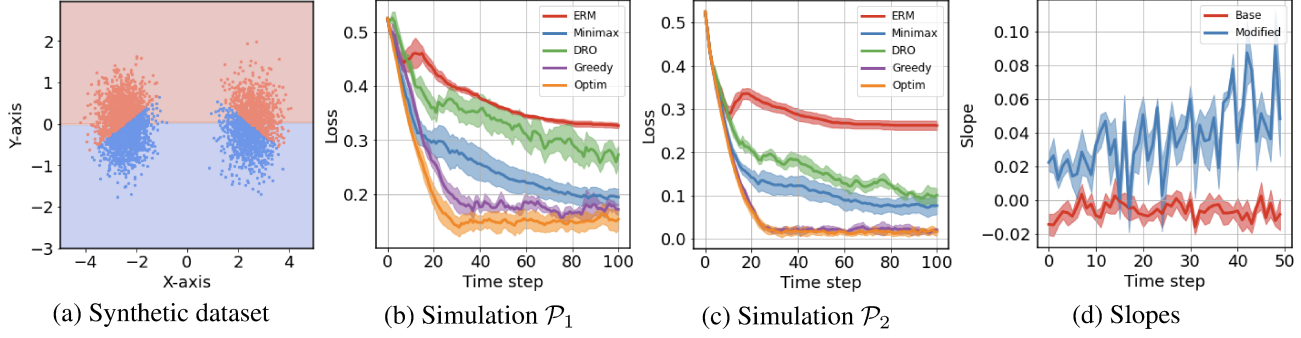


Figure 1. (a) illustrates the synthetic dataset with a fair model decision boundary. (b) and (c) plot the binary cross-entropy losses of population densities resulting from ERM, Minimax, DRO, Greedy, and the proposed Optim methods in simulation environments $\mathcal{P}_1$ and $\mathcal{P}_2$ respectively. (d) plots model slopes resulting from the Optim method in the base and modified environments.

of both the birth and survival functions. It then generates a sequence of models utilizing the method of successive approximation, as detailed in Algorithm 1.

Figure 1 (b) and (c) demonstrate the results simulated from $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. As observed, in both simulation environments, ERM results in comparably high losses at all time steps due to the low population densities in the minority group. For instance, at $t = 100$, ERM results in $\lambda_T^1 = 1$ and $\lambda_T^2 = 0.53$ with a loss of 0.32. For DRO, we utilize the initial minority population density to compute the DRO loss. This approach yields conservative models at all time steps, leading to low densities accompanied by high losses. Minimax is capable of mitigating representation disparity by minimizing the worst-case loss among both groups, thereby resulting in a more balanced terminal state with $\lambda_T^1 = 0.9$ and $\lambda_T^2 = 0.76$. It is noteworthy that the aforementioned methods do not consider the underlying dynamics. Conversely, the optimal and greedy control methods implement T -step and 1-step planning respectively, leveraging the estimated underlying dynamics. When these dynamics significantly impact user retention, the T -step planning demonstrates superior performance compared to the single-step planning, as depicted in Fig.1 (b) under simulation environment $\mathcal{P}_1$. However, simulation environment $\mathcal{P}_2$ stipulates user churn when 3 wrong predictions have accrued, this leads to a system with smoother population dynamics, where T -step planning performs comparably to the single-step planning method, as shown in Fig.1 (c).

The advantage of trajectory planning that acknowledges the underlying dynamics is explored herein. The idea is to interpret the evolution of the population retention system at certain time steps and observe corresponding adjustments made in the optimal control solution. The optimal control method makes performance tradeoffs from the majority group to balance the population densities of both groups at the terminal step. We manually introduce a substantial quantity of users into the minority demographic group at

$t = 50$. Consequently, it is expected that the optimal control method would make fewer tradeoffs and adjust its decision boundaries accordingly at earlier time steps (e.g., $t < 50$). Referring to Fig.1 (a), we consider a linear classifier where a positive (resp. negative) slope indicates a model favoring the majority demographic (resp. minority) group. Figure 1 (d) illustrates the slopes of the model decision boundary at $t \in [0, 50]$. As observed, the introduction of additional users to the minority group at $t = 50$ enables the optimal control solution to make less performance tradeoff against the majority group due to the increased population density at a later time step. This adjustment cannot be accomplished with greedy control-based methods.

5.3. Modeling with Adult Income and COMPAS Recidivism Racial Bias.

We explore two real-world datasets: Adult Income (Adult) (Kohavi et al., 1996) and COMPAS (Barenstein, 2019). The Adult dataset describes an individual’s annual income, influenced by factors such as gender, race, age, and education level, while COMPAS is a widely-used commercial tool that assists judges and parole officers in assessing the likelihood of a criminal defendant reoffending. In both datasets, gender attributes are used to differentiate demographic groups. The population retention system, as defined in Eq. (1), is applied to these static datasets to simulate population dynamics. A total of $N = 3000$ samples are randomly selected from each demographic group, setting initial population densities at $\lambda_0^1 = 0.7$ and $\lambda_0^2 = 0.5$ for the majority and minority groups respectively. Fig. 2 (a) and (b) illustrate the trajectories of the binary cross-entropy loss in the environments $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. The optimal control solution (Optim) surpasses other baselines, achieving terminal states with $\lambda_T^1 = 0.93$ and $\lambda_T^2 = 0.78$ in $\mathcal{P}_1$, and $\lambda_T^1 = 1$ and $\lambda_T^2 = 0.81$ in $\mathcal{P}_2$. Fig. 2 (c) and (d) present the results obtained from the COMPAS dataset. In both instances, due to the relatively smooth dynamics inherent in the COMPAS dataset, single-step planning performs on par with the

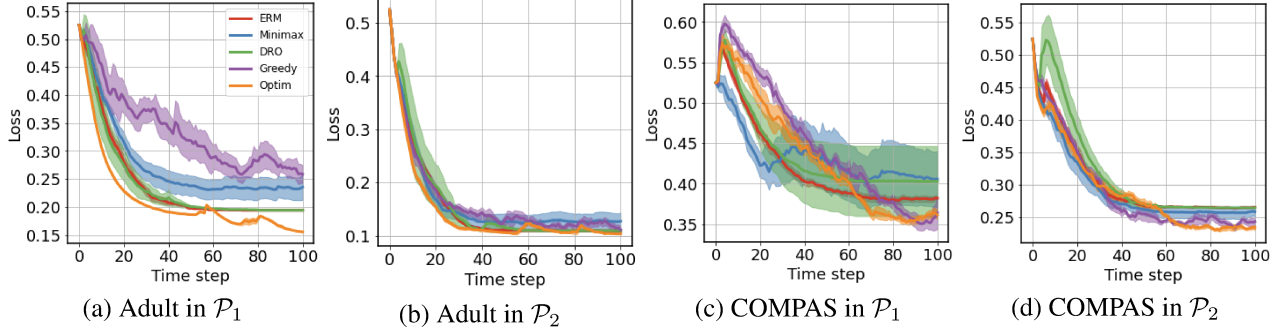


Figure 2. (a) and (b) depict the trajectories of losses $\Psi(\lambda_t, 1)$ resulting from simulation environments $\mathcal{P}_1$ and $\mathcal{P}_2$ respectively, utilizing the Adult dataset. Similarly, the outcomes from $\mathcal{P}_1$ and $\mathcal{P}_2$ using the COMPAS dataset are illustrated in figures (c) and (d) respectively.

optimal control solution.

6. Limitations and Future Works

In this section, we provide a discussion of the limitations of the proposed framework, thereby highlighting potential areas for improvement and development.

Constraints on intermediate population densities: In the context of fairness, asymptotic behavior plays a crucial role in defining long-term population densities. However, exclusive dependence on asymptotic fairness might fall short in managing representation disparities during intermediate time steps, potentially resulting in unsatisfying user experiences. To mitigate this limitation, our future work aims to incorporate an additional running loss into the framework.

Surrogate retention system: In real-world scenarios, user churn behavior is influenced by a multitude of factors, including financial background and health conditions. These factors are not static but vary over time, which requires the adoption of a dynamic modeling approach. Existing literature in the field of evolutionary population dynamics often employs a canonical equation, known as Darwinian dynamics, to describe the evolution of individuals' traits (Cushing, 2019). However, to achieve a more accurate representation of population dynamics, it is essential to incorporate individual characteristics into the modeling framework.

Computational efficiency: The current framework addresses the control problem using the PMP, which involves solving an optimization problem. However, this approach becomes inefficient when the algorithm is deployed in real-world services that require frequent updates. Consequently, our next objective is to derive an accurate closed-form solution that maintains the desired fairness objectives. By refining the closed-form solution, we aim to develop a more efficient and effective algorithm that can be seamlessly integrated into services requiring frequent updates, while still

preserving the desired level of fairness.

7. Conclusion

With the growing integration of machine learning systems into real-world applications, studying long-term fairness problems holds substantial significance. In this work, we have presented a framework to investigate fairness in a dynamic setting, where user participation is conditioned on model performance. We have defined asymptotic fairness to evaluate fairness within this dynamic setting and formulated it as an optimal control problem. To accomplish this, we have constructed a surrogate retention system that provides estimates of the underlying environment. The proposed optimal control solution has been demonstrated to be effective through simulations. Future work will focus on designing more sophisticated surrogate retention systems capable of managing complex environments, potentially involving real-world human users. Additionally, Another potential avenue is to seek to theoretically comprehend the working principle of the optimal control solution, emphasizing the distinction from greedy control-based methods. We aim to highlight how the underlying dynamics influence the determination of optimal control solutions in the context of fairness in a non-stationary environment.

References

- Barenstein, M. Propublica's compas data revisited. *arXiv preprint arXiv:1906.04711*, 2019.
- Bellman, R. On the theory of dynamic programming. *Proceedings of the National Academy of Sciences of the United States of America*, 38(8):716, 1952.
- Chen, Z., Li, Q., and Zhang, Z. Towards robust neural networks via close-loop control. In *International Conference on Learning Representations*, 2021.
- Chen, Z., Li, Q., and Zhang, Z. Self-healing robust neural networks via closed-loop control. *Journal of Machine Learning Research*, 23(319):1–54, 2022. URL <http://jmlr.org/papers/v23/22-0529.html>.
- Cushing, J. Difference equations as models of evolutionary population dynamics. *Journal of Biological Dynamics*, 13(1):103–127, 2019.
- Diana, E., Gill, W., Kearns, M., Kenthapadi, K., and Roth, A. Minimax group fairness: Algorithms and experiments. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 66–76, 2021.
- Duchi, J. C., Hashimoto, T., and Namkoong, H. Distributionally robust losses against mixture covariate shifts. *Under review*, 2, 2019.
- E. A proposal on machine learning via dynamical systems. *Communications in Mathematics and Statistics*, 5(1):1–11, 2017.
- E, W., Han, J., and Li, Q. A mean-field optimal control formulation of deep learning. 6(1):10. ISSN 2522-0144.
- Feldman, M., Friedler, S. A., Moeller, J., Scheidegger, C., and Venkatasubramanian, S. Certifying and removing disparate impact. In *proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 259–268, 2015.
- Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., and Bouchachia, A. A survey on concept drift adaptation. *ACM computing surveys (CSUR)*, 46(4):1–37, 2014.
- Haber, E. and Ruthotto, L. Stable architectures for deep neural networks. *Inverse Problems*, 34(1):014004, 2017.
- Hardt, M., Price, E., and Srebro, N. Equality of opportunity in supervised learning. *Advances in neural information processing systems*, 29, 2016.
- Harwell, D. Amazon's alexa and google home show accent bias, with chinese and spanish hardest to understand. 2018.
- Hashimoto, T., Srivastava, M., Namkoong, H., and Liang, P. Fairness without demographics in repeated loss minimization. In *International Conference on Machine Learning*, pp. 1929–1938. PMLR, 2018.
- Kirk, D. E. *Optimal control theory: an introduction*. Springer, 1970.
- Kohavi, R. et al. Scaling up the accuracy of naive-bayes classifiers: A decision-tree hybrid. In *Kdd*, volume 96, pp. 202–207, 1996.
- Li, Q. and Hao, S. An optimal control approach to deep learning and applications to discrete-weight neural networks. In *International Conference on Machine Learning*, pp. 2985–2994. PMLR, 2018.
- Li, Q., Chen, L., Tai, C., and E, W. Maximum principle based algorithms for deep learning. *Journal of Machine Learning Research*, 18(165):1–29, 2018. URL <http://jmlr.org/papers/v18/17-653.html>.
- Liu, G.-H. and Theodorou, E. A. Deep learning theory review: An optimal control and dynamical systems perspective. *arXiv preprint arXiv:1908.10920*, 2019.
- Liu, G.-H., Chen, T., and Theodorou, E. A. Differential dynamic programming neural optimizer. *arXiv preprint arXiv:2002.08809*, 2020a.
- Liu, G.-H., Chen, T., and Theodorou, E. A. A differential game theoretic neural optimizer for training residual networks. *arXiv preprint arXiv:2007.08880*, 2020b.
- Liu, L. T., Dean, S., Rolf, E., Simchowitz, M., and Hardt, M. Delayed impact of fair machine learning. In *International Conference on Machine Learning*, pp. 3150–3158. PMLR, 2018.
- Pontryagin, L. S. *Mathematical theory of optimal processes*. CRC press, 1987.
- Schlimmer, J. C. and Granger, R. H. Incremental learning from noisy data. *Machine learning*, 1(3):317–354, 1986.
- Vincent, T. L. and Brown, J. S. *Evolutionary game theory, natural selection, and Darwinian dynamics*. Cambridge University Press, 2005.
- Widmer, G. and Kubat, M. Learning in the presence of concept drift and hidden contexts. *Machine learning*, 23(1):69–101, 1996.
- Zhang, X., Khaliligarekani, M., Tekin, C., et al. Group retention when using machine learning in sequential decision making: the interplay between user dynamics and fairness. *Advances in Neural Information Processing Systems*, 32, 2019.

Zhang, X., Tu, R., Liu, Y., Liu, M., Kjellstrom, H., Zhang,
K., and Zhang, C. How do fair decisions fare in long-
term qualification? *Advances in Neural Information
Processing Systems*, 33:18457–18469, 2020.