

# A Bias-Aware Deep Learning Framework for Hierarchical Microcredential Classification

Mohammad Arif Ul Alam  
LivedX Inc.  
University of Massachusetts  
Lowell  
malam@cs.uml.edu

Geeta Verma  
LivedX Inc.  
University of Colorado Denver  
geeta.verma@ucdenver.edu

Eumie Jhong  
LivedX Inc.  
eumie@livedx.com

Justin Barber  
LivedX Inc.  
Pearson  
justin.o.barber@gmail.com

Ashis Kumer Biswas  
LivedX Inc.  
University of Colorado Denver  
ashis.biswas@ucdenver.edu

## ABSTRACT

The growing demand for microcredentials in education and workforce development necessitates scalable, accurate, and fair assessment systems for both soft and hard skills based on students' lived experience narratives. Existing approaches struggle with the complexities of hierarchical credentialing and the mitigation of algorithmic bias related to gender and ethnicity. In this paper, we propose a novel deep learning framework that integrates hierarchical classification based on dynamic thresholding with a dual deep Q network dueling (DDQN dueling) for bias mitigation. Our method improves predictive performance at all three levels of microcredential classification, achieving an increase in 7% sensitivity and an improvement in 6% specificity over baseline models<sup>1</sup>. Furthermore, our framework significantly improves fairness by reducing gender and ethnicity bias, as measured by equalized odds, by over 20% compared to conventional approaches. Extensive evaluations on a dataset of 3,000 student narratives demonstrate a 12% improvement in the F1 score and a 5% increase in AUROC relative to existing methods. These results underscore the effectiveness of our approach in advancing both hierarchical classification accuracy and fairness in real-world educational applications.

## Keywords

Hierarchical microcredentialing, Reinforcement learning, Bias mitigation, Student narratives, Deep learning embeddings

## 1. INTRODUCTION

In recent years, microcredentials have become a vital component of both higher education and workforce development, providing learners with a way to demonstrate hard and soft skills that traditional degree programs often overlook. The global market for soft-

skills assessment tools is projected to reach \$1.2 billion by 2028, with institutions and employers increasingly recognizing the value of these credentials for student and employee development [9]. In academia, microcredentials help bridge the gap between academic outcomes and job-ready skills, offering students tangible recognition of competencies that can enhance their employability [5]. Currently, 96% of the U.S. workforce is actively seeking opportunities for career advancement, underscoring the importance of scalable and structured microcredential systems [19].

Traditional methods of awarding academic credits through microcredentials are largely manual, requiring human annotators to review students' portfolios and lived experiences. This process is time-consuming, labor-intensive, and subject to inconsistencies due to subjective judgment [11]. Furthermore, the complexity of hierarchical microcredential structures where credentials range from broad foundational skills to specialized competencies presents significant challenges for manual annotation. The hierarchical structure not only requires annotators to understand the relationships between different levels of credentials but also to classify and assign them accurately, which becomes infeasible at scale. As a result, natural language processing (NLP) and deep learning techniques are being increasingly used to automate microcredential classification, enabling more efficient and consistent assessment of unstructured data such as student narratives [26].

Several studies have explored the use of automated tools for classifying microcredentials from student submissions. Initial efforts have typically focused on using rule-based systems or simple NLP techniques to extract keywords and assign credentials based on predefined rubrics [27]. However, these methods often fail to capture the rich, unstructured narratives found in students' lived experiences, which contain important indicators of soft skills such as leadership, collaboration, and problem-solving. Extracting meaningful insights from such narratives requires more sophisticated models capable of handling context-dependent information. Moreover, developing a large, annotated dataset of lived experience narratives is a significant challenge, given the variability in how students express their experiences. Advanced deep learning models, particularly those that can handle hierarchical structures, are essential for automating the extraction and classification of microcredentials from these narratives [25].

A significant challenge in developing automated microcredential-

<sup>1</sup>A preliminary analysis of raw narratives without hierarchical credential labels appears in [1, 3].

Mohammad Arif Ul Alam, Geeta Verma, Eumie Jhong, Justin Barber, and Ashis Kumer Biswas. A Bias-Aware Deep Learning Framework for Hierarchical Microcredential Classification. In Caitlin Mills, Giora Alexandron, Davide Taibi, Giosuè Lo Bosco, and Luc Paquette (eds.) Proceedings of the 18th International Conference on Educational Data Mining, Palermo, Italy, July, 2025, pp. 664–672. International Educational Data Mining Society (2025).

© 2025 Copyright is held by the author(s). This work is distributed under the Creative Commons Attribution NonCommercial NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.  
<https://doi.org/10.5281/zenodo.15870151>

ing systems is the risk of algorithmic bias, especially in terms of gender and ethnicity. Biases present in the training data can lead machine learning models to systematically misclassify or undervalue the competencies of underrepresented groups [2]. For example, differences in how men and women, or individuals from different ethnic backgrounds, articulate their experiences may result in biased outcomes, with certain groups receiving fewer or less favorable microcredentials. Addressing this issue is critical, as biased systems could exacerbate existing inequalities in education and employment opportunities [20].

In this paper, we address the key challenges in automatic microcredential classification by proposing a novel framework that leverages deep learning techniques to automate the process in a fair and accurate manner. Our contributions are as follows:

- We design and annotate a 3-tier hierarchical structure for microcredentials based on students’ lived experience narratives, capturing both general and specialized competencies. This design allows for a more structured and scalable approach to microcredentialing, providing flexibility and precision across various levels of skills.
- We introduce a dynamic thresholding-based hierarchical classification model that adapts to different levels of microcredentials, enabling more accurate predictions even in data-scarce categories. Our approach ensures that each level of the hierarchy is treated appropriately, improving overall classification performance.
- To address biases related to gender and ethnicity, we implement a Dueling Double-Deep Q-Network (Dueling DDQN) for reinforcement learning-based bias mitigation. This model learns optimal decision-making policies that balance fairness and performance, ensuring equitable outcomes across demographic groups while maintaining classification accuracy.

Our two previous works analyzed the same LivedX narrative datasets: one leveraged large language model augmentation to predict social determinants of mental health from student essays[1], and another investigated human–AI annotation workflows to improve ethical outcomes[3]. Neither addressed hierarchical microcredential inference, dynamic threshold calibration, nor fairness aware learning. The present work closes these gaps by (i) introducing a novel three tier micro credential taxonomy with newly curated labels, (ii) proposing an attention guided hierarchical classifier that fuses regression based count estimations with dynamic thresholding, and (iii) deploying a Dueling Double Deep Q Network that reduces gender and ethnicity based equalised odds disparities by up to 35%.

## 2. RELATED WORKS

Research on microcredentials has gained significant attention in recent years as a growing number of academic institutions and industries seek scalable and reliable methods for assessing and awarding credentials [23, 14].

### 2.1 Microcredential Awarding Methods

Microcredentialing has evolved significantly, with initial systems relying heavily on manual and rubric-based assessment methods [14, 27]. Early frameworks such as those presented in [27] utilized predefined rubrics to classify credentials, often focusing on structured datasets. The introduction of machine learning in microcredential awarding, such as the work by [15, 28], shifted this process to a more scalable approach. However, these methods were largely focused on structured or semi-structured inputs, and they lacked the

ability to handle more complex, unstructured lived experiences [10]. Some approaches, like [18], explored lived experience narratives to infer soft skills, but these studies were limited in scope and failed to implement a hierarchical structure. Recently, microcredential frameworks have explored hierarchical categorization, allowing a more fine-grained distinction between skills across levels [23, 14]. Yet, hierarchical microcredential systems remain underdeveloped, especially in automatically processing unstructured text data, where NLP and deep learning models hold potential but remain underexplored [22]. We have [1, 3] analyzed the same narrative corpus for mental health features and annotation quality; did not attempt hierarchical credential inference nor fairness mitigation.

### 2.2 Hierarchical Classification from Texts

Hierarchical classification of text data is an area that has seen rapid development in recent years. Traditional classification models often struggle to maintain accuracy when tasked with distinguishing between various levels of a hierarchy, as errors in top-level categories can propagate to lower levels, compounding misclassification rates [32]. To address this issue, several works have proposed enhanced hierarchical classification techniques, such as the use of hierarchical attention networks [35], which attempt to weigh different parts of a narrative more heavily depending on their importance at different levels of the hierarchy. Recent work by [32] introduced dynamic models that adjust to class imbalance issues inherent in hierarchical classification. However, these models tend to rely on static thresholds, which limit their flexibility. Our proposed dynamic thresholding approach builds on these efforts by allowing for real-time adjustments to classification boundaries, improving both precision and recall across levels, particularly in data-scarce classes. This technique addresses the need for adaptable models that can better manage the complexity of hierarchical microcredential classification from textual narratives.

### 2.3 Bias Mitigation in Deep Learning and NLP

Bias in machine learning models, particularly those used in natural language processing (NLP), has emerged as a critical issue, especially when these models are deployed in sensitive applications like education and workforce development [15, 28]. Several studies have explored techniques to mitigate bias, such as debiasing word embeddings [4] and using adversarial training to minimize group disparities [36]. While these approaches have shown success in reducing biases related to gender and ethnicity, they often lack the flexibility to handle bias dynamically across hierarchical structures [8]. Reinforcement learning has been increasingly applied to mitigate bias in decision-making models [16, 29, 7, 33, 24], with notable approaches such as the Dueling Double-Deep Q-Network (DDQN) framework showing promise in areas like resource allocation and policy optimization [34, 33, 7, 29]. Our work extends this by applying DDQN to the NLP domain for bias mitigation in hierarchical microcredential classification. By integrating fairness constraints into the reward structure of the DDQN, we ensure that the system learns to balance accuracy with equitable outcomes, addressing some of the limitations of static bias mitigation techniques found in earlier studies.

## 3. HIERARCHICAL MICROCREDENTIALS DATA COLLECTION AND ANNOTATION

### 3.1 Data Collection

Detailed characteristics of the raw narrative dataset were first reported in our earlier studies [1, 3]; we reuse the corpus here but supply new hierarchical skill labels and fairness metadata. Data for

this project were collected through an online platform, publicly accessible web-based interface called 'LivedX'<sup>2</sup>. Students were systematically guided to document their lived experiences on the platform. The platform framework, drawing upon phenomenological perspectives [17] and the Funds of Knowledge framework [21], was used to interpret the multifaceted nature of these experiences. This framework credentialized the embedded power skills in students' narratives, issuing microcredentials, culminating in a comprehensive portfolio showcasing the diverse skill sets and competencies students had developed. This empowered individuals to articulate their skills, enhancing their capacity to navigate a skill-dependent landscape.

All data collection and analyses were conducted under LivedX's independent Institutional Review Board (IRB) approval, which authorizes the use of de identified, aggregated student narratives for algorithm refinement and other research purposes. Upon account creation, students electronically accept a Terms of Use and Research Consent clause reviewed by the IRB that explicitly permits such aggregated reuse. Narratives are automatically scrubbed of direct identifiers before storage, and only population level statistics are ever released; no individual level data leave the secure research server. No monetary incentives were offered beyond normal platform access. Students did not have access to the full three tier micro credential taxonomy while composing their narratives. At sign in they were only shown seven broad competency bands (e.g. Communication, Collaboration); the fine grained labels used for model training remained hidden to prevent response priming. Table 1 provides an overview of the hierarchical structure, including representative examples for each level.

## 3.2 Human Annotation Process

A team of six human annotators was involved in labeling the dataset, ensuring consistency and accuracy in microcredential classification. The annotators were selected based on their backgrounds in education, social-emotional learning, and prior experience in qualitative data analysis. Each annotator had prior experience in evaluating students' competencies, with at least two having prior teaching or assessment experience in education or workforce training programs. Additionally, two annotators had domain expertise in social-emotional learning and qualitative data analysis, while others had experience in human-centered AI research and narrative assessment. To enhance annotation consistency and reduce subjective biases, all annotators underwent a structured training process. Initially, they labeled a subset of narratives independently based on a predefined hierarchical rubric. After this initial phase, they received training from social-emotional learning experts and were provided with example annotations to ensure alignment with best practices. The annotation process was iterative, with weekly calibration sessions where discrepancies were discussed, and annotation guidelines were refined accordingly. Diversity in annotation was also a key consideration, ensuring representation from different demographic backgrounds. This helped mitigate unintended biases that could arise from a homogeneous annotator group. Interrater reliability (IRR) was calculated to assess the validity and robustness of the qualitative data analysis. IRR measures the agreement among independent coders categorizing qualitative data [31]. Based on power calculations, six annotators were required to ensure sufficient power for comparing pre- and post-training results. Each rater was assigned 42-45 items to ensure an adequate sample size for statistical analysis, as confirmed by G\*Power software.

<sup>2</sup><https://livedx.com/>

Table 2 presents pre- and post-training IRR results. Post-training assessments indicated a significant improvement in interrater reliability (IRR), increasing from 0.46 in pre-training annotations to 0.83 post-training. Training improved the accuracy and consistency of annotations, reflected in the enhanced intra-class correlation coefficient (ICC) scores, rising from 0.46 pre-training to 0.83 post-training. A complete list and example of hierarchical microcredentials is available in the supplementary materials.

## 3.3 Rationale for Annotators and Submissions

The decision to use six annotators and target 42-45 submissions follows established guidelines for achieving reliable IRR. According to [31], a minimum of two raters is required for 43 pieces of information to ensure adequate power for obtaining an ICC above 0.70. Given the comparison between pre- and post-training results, four raters were necessary to ensure statistical significance when comparing frameworks. Expanding to six annotators enabled t-tests and comparisons with greater statistical power, as suggested by G\*Power software. Leveraging these annotated data, we developed and refined a machine learning algorithm. The combination of human expertise and automated computational systems allowed the platform to issue microcredentials with high accuracy and consistency. The annotation process was critical in ensuring that the machine learning models were trained on high-quality datasets.

## 3.4 Illustrative Examples & Educational Significance of the Hierarchy

**Alignment with recognized frameworks:** Our 3 tier taxonomy mirrors the stackable approach recommended by Common Micro credential Framework (CMF), which specifies that short, workforce relevant learning blocks should be organized in nested levels that can be 'stacked' towards larger awards [13]. Similar multi level structures underpin Europe's policy blueprint for micro credentials [12] and UNESCO's 2023 global guidelines, which emphasise transparency of learning outcomes at each depth of mastery [30]. Outside the policy arena, institutional playbooks such as Northeastern University's 4 level badging model likewise stress a progression from broad competencies to granular skills evidence [6]. By mapping lived experience narratives to Communication → Communication Qualities → Audience Centered Delivery, our hierarchy operationalizes these principles in the higher education setting.

**Quantitative insight:** Across the 3,000 annotated narratives the mean micro credential count is  $2.4 \pm 0.9$  per submission. Level 1 classes are well balanced (Social Emotional Learning 28%, Academic/Professional 34%, Collaboration 23%, Communication 15%), suggesting the rubric captures the breadth of students' self reported growth. The Gini index at Level 3 is 0.27, indicating that no single fine grained credential dominates, a desirable trait for formative, learner centered assessment.

## 4. DYNAMIC THRESHOLDING BASED HIERARCHICAL PREDICTION WITH ATTENTION

The task of hierarchical microcredential prediction can be challenging due to the multi-level nature of the labels and the need to predict both the classes and the number of microcredentials assigned to each class. In this section, we propose a novel approach that combines regression and classification at each level of the hierarchy with an attention mechanism. Moreover, we introduce a dynamic thresholding mechanism for level 3 microcredential classification,

Table 1: Structured Table for Social Emotional Learning, Academic and Professional Skills, and Communication

Level I	Level II	Level III
Social Emotional Learning	Recognize/manage emotions, thoughts, behaviors; Perspective taking	Recognizes others' emotions, thoughts, behaviors, body sensations; Observes situations; Understands differing viewpoints
Academic & Professional Skills	Responsible behavior; Leadership	Fulfills expectations (job/school/home); Plans and organizes; Manages human resources
Collaboration	Working with others	Respects differences/opinions; Builds relationships with diverse individuals; Negotiates conflicts; Demonstrates/understands teamwork (e.g., school project)
Communication	Communication qualities	Communicates with groups, audiences, stakeholders

Table 2: Interrater Reliability Pre- and Post-Training Results

Measure	Pre	Post
Avg Correct Answers	59%	83%
Avg Agreement	59%	70%
Items Matched Correct & Agreement	21 (47%)	22 (50%)
ICC	0.46	0.83

Table 3: Illustrative student narratives mapped to the three tier microcredential hierarchy.

<b>Narrative A (STEM Hackathon Lead).</b> "I organised a 24 hour campus hackathon for 150 peers, secured \$5,000 sponsorship, and mediated disputes between teams during judging." Level 1: COLLABORATION Level 2: Leadership & Project Management Level 3: Resource Mobilisation, Conflict Negotiation, Agile Scheduling
<b>Narrative B (Community Health Volunteer).</b> "Each weekend I translate discharge instructions into Spanish for rural clinics and teach patients how to schedule follow up visits online." Level 1: COMMUNICATION Level 2: Communication Qualities (Clarity, Cultural Competence) Level 3: Cross lingual Mediation, Digital Literacy Coaching, Audience Centred Delivery

which aligns with the regression predictions of higher-level classes, ensuring coherence across all levels.

## 4.1 Problem Formulation

We address the hierarchical microcredential prediction problem where students' lived experience narratives are labeled in a hierarchical structure across three levels:

**Level 1:** 8 top-level categories of microcredentials.

**Level 2:** 32 subcategories within each of the Level 1 categories.

**Level 3:** 152 granular microcredential categories, which are the most detailed level of the hierarchy.

The goal is to predict multi-label outputs for each of the three levels, where the labels at higher levels constrain the labels at lower levels. In particular, the sum of predicted Level 3 microcredentials for a given Level 1 or Level 2 class should align with the predicted number of microcredentials at those higher levels.

## 4.2 Preliminaries

Let the set of student experiences be denoted by  $\mathcal{E} = \{e_1, e_2, \dots, e_n\}$ , where each experience is described as a free-text narrative. For each experience, we aim to predict a multi-label classification vector for Level 1, Level 2, and Level 3 microcredentials:

$$\mathbf{y}_1 = [y_{1,1}, y_{1,2}, \dots, y_{1,8}], \quad (1)$$

$$\mathbf{y}_2 = [y_{2,1}, y_{2,2}, \dots, y_{2,32}], \quad (2)$$

$$\mathbf{y}_3 = [y_{3,1}, y_{3,2}, \dots, y_{3,152}], \quad (3)$$

where each  $y_{i,j}$  is a binary label indicating whether a specific microcredential class is present in experience. For each experience, we predict number of microcredentials assigned to each class, which introduces a regression problem.

## 4.3 Hierarchical Regression and Classification with Attention

To tackle this problem, we develop a multi-stage model that leverages attention mechanisms to perform both regression and classification at each hierarchical level. The framework is implemented in three stages:

**Level 1: Regression and Classification:** At first level, we aim to predict both the presence of the eight top-level microcredential categories and number of Level 3 microcredentials within each category. We employ a combined regression and classification model with an attention layer:

$$\hat{\mathbf{y}}_1^{\text{reg}} = \text{Regressor}(\mathbf{h}_1), \quad (4)$$

$$\hat{\mathbf{y}}_1^{\text{class}} = \sigma(\text{Classifier}(\mathbf{h}_1) \cdot \mathbf{A}_1), \quad (5)$$

where  $\mathbf{h}_1$  represents the combined embeddings generated from a sentence transformer,  $\mathbf{A}_1$  is an attention matrix, and  $\sigma$  is the sigmoid activation function. The regression component predicts the count of Level 3 microcredentials for each Level 1 category, and the classification component provides multi-label outputs for the presence or absence of the top-level categories. The loss function for Level 1 combines the mean squared error (MSE) for regression and binary cross-entropy (BCE) for classification:

$$\mathcal{L}_1 = \text{MSE}(\hat{\mathbf{y}}_1^{\text{reg}}, \mathbf{y}_1^{\text{reg}}) + \text{BCE}(\hat{\mathbf{y}}_1^{\text{class}}, \mathbf{y}_1^{\text{class}}). \quad (6)$$

**Level 2: Regression and Classification:** At Level 2, the same architecture is applied but extended to 32 classes. The predicted regression outputs for Level 2 categories are constrained by Level 1 regression outputs. The model performs regression to predict number of Level 3 microcredentials for each of Level 2 categories and multi-label classification for presence of those categories:

$$\hat{\mathbf{y}}_2^{\text{reg}} = \text{Regressor}(\mathbf{h}_2), \quad (7)$$

$$\hat{\mathbf{y}}_2^{\text{class}} = \sigma(\text{Classifier}(\mathbf{h}_2) \cdot \mathbf{A}_2), \quad (8)$$

where  $\mathbf{h}_2$  is the embedding at Level 2 and  $\mathbf{A}_2$  is the attention matrix for Level 2. The loss function is analogous to Level 1, combining MSE and BCE losses.

**Level 3: Classification with Attention:** At the most granular level, the model is designed to predict the presence of 152 microcredential classes (Level 3) using only classification. The attention mechanism plays a crucial role in weighting the contributions of different features to the final predictions:

$$\hat{\mathbf{y}}_3^{\text{class}} = \sigma(\text{Classifier}(\mathbf{h}_3) \cdot \mathbf{A}_3). \quad (9)$$

Here,  $\mathbf{h}_3$  is the embedding generated for Level 3, and  $\mathbf{A}_3$  is the attention matrix that focuses on the most relevant features with loss:

$$\mathcal{L}_3 = \text{BCE}(\hat{\mathbf{y}}_3^{\text{class}}, \mathbf{y}_3^{\text{class}}). \quad (10)$$

#### 4.4 Dynamic Threshold Selection for Lowest Level Classification

The final step in our framework is to dynamically select the threshold for Level 3 classification based on the outputs of the regression models at Levels 1 and 2. The dynamic thresholding mechanism ensures that the number of Level 3 microcredentials predicted for each Level 1 and Level 2 category matches the regression outputs from those levels.

Let  $\mathbf{T}_3$  represent the set of candidate thresholds. For each threshold  $t \in \mathbf{T}_3$ , we predict the Level 3 microcredentials:

$$\hat{\mathbf{y}}_3^{\text{class}}(t) = \mathbb{I}(\hat{\mathbf{y}}_3^{\text{class}} > t), \quad (11)$$

where  $\mathbb{I}$  is the indicator function. For each threshold, we count the predicted Level 3 microcredentials for each Level 1 and Level 2 category and compute the Euclidean distance between these counts and the regression outputs from the previous stages:

$$\text{dist}(t) = \|\hat{\mathbf{y}}_3^{\text{class}}(t) - \hat{\mathbf{y}}_1^{\text{reg}}\|_2 + \|\hat{\mathbf{y}}_3^{\text{class}}(t) - \hat{\mathbf{y}}_2^{\text{reg}}\|_2. \quad (12)$$

The optimal threshold  $t^*$  is chosen on minimal total distance:

$$t^* = \arg \min_{t \in \mathbf{T}_3} \text{dist}(t). \quad (13)$$

This dynamic thresholding process ensures coherence between the regression and classification models across all hierarchical levels.

#### 4.5 Attention Mechanism and Final Prediction

The attention mechanism at each level plays a crucial role in enhancing the performance of both the regression and classification tasks. By applying attention weights to the predicted scores, the model is able to focus on the most relevant features for each hierarchical level. The final output for each experience is a set of Level 3 microcredentials, adjusted dynamically based on the regression outputs at Levels 1 and 2.

The complete model thus provides a robust solution to the hierarchical microcredential prediction problem by combining attention-based regression, multi-label classification, and dynamic thresholding to ensure consistency across all hierarchical levels.

---

##### Algorithm 1 Dueling DDQN for Bias Mitigation

---

```

1: Initialize replay memory  $\mathcal{D}$  to capacity  $N$ 
2: Initialize main network  $Q(s, a; \theta)$  and target network  $Q'(s, a; \theta^-)$  with random weights
3: for episode = 1 to  $M$  do
4:   Initialize state  $s_1$ 
5:   for  $t = 1$  to  $T$  do
6:     With probability  $\epsilon$ , select a random action  $a_t$ , otherwise
       select  $a_t = \arg \max_a Q(s_t, a; \theta)$ 
7:     Execute action  $a_t$ , observe reward  $r_t$  and next state  $s_{t+1}$ 
8:     Store transition  $(s_t, a_t, r_t, s_{t+1})$  in  $\mathcal{D}$ 
9:     Sample a mini-batch of transitions  $(s_j, a_j, r_j, s_{j+1})$  from  $\mathcal{D}$ 
10:    Set  $y_j = r_j + \gamma Q'(s_{j+1}, \arg \max_a Q(s_{j+1}, a; \theta))$ 
11:    Perform a gradient descent step on  $(y_j - Q(s_j, a_j; \theta))^2$ 
12:    Every  $C$  steps, update target network:  $\theta^- \leftarrow \theta$ 
13:  end for
14: end for

```

---

#### 5. BIAS MITIGATION USING REINFORCEMENT LEARNING

Our goal is to mitigate bias in predictive modeling by using reinforcement learning to address disparities in prediction outcomes across different subgroups, specifically for gender and ethnicity. Formally, let  $\mathbf{X}$  be the set of features,  $\mathbf{Y} \in \{0, 1\}$  be the binary class labels, and  $\mathbf{Z} \in \{0, 1\}$  represent a protected attribute, such as gender or ethnicity, where  $Z = 1$  signifies the minority group. The aim is to develop a classifier  $\hat{Y}$  that satisfies fairness criteria, such as equalized odds, which ensures that the true positive rate (TPR) and false positive rate (FPR) are equal for both protected and unprotected groups:

$$P(\hat{Y} = 1 \mid Y = 1, Z = 0) = P(\hat{Y} = 1 \mid Y = 1, Z = 1)$$

$$P(\hat{Y} = 1 \mid Y = 0, Z = 0) = P(\hat{Y} = 1 \mid Y = 0, Z = 1)$$

This ensures that predictions do not disproportionately affect one group over another.

##### 5.1 Dueling Double-Deep Q-Network

We propose a Dueling Double-Deep Q-Network (Dueling DDQN) to address the bias in microcredential predictions, specifically with respect to gender and ethnicity as protected variables. Dueling DDQN separates the action-value function into two streams: one for the state-value and the other for the action advantages, as defined by the following equations:

$$Q(s, a; \theta, \alpha, \beta) = V(s; \beta) + \left( A(s, a; \alpha) - \frac{1}{|A|} \sum_{a'} A(s, a'; \alpha) \right)$$

Where  $V(s; \beta)$  is the state value,  $A(s, a; \alpha)$  is the advantage function, and  $Q(s, a)$  represents the Q-value for state-action pairs.

This formulation helps the network better distinguish between the importance of a state and the advantage of a particular action in mitigating bias, without having to estimate action values for every possible action.

##### 5.2 Reward Design for Bias Mitigation

To ensure fairness in the predictions across protected groups, the reward function in our RL framework is designed to encourage correct classification of both majority and minority classes, while providing a bias-sensitive penalty for misclassifications. We define a reward  $R(s_t, a_t)$  that takes into account both the classification accuracy and fairness across protected attributes:

$$R(s_t, a_t) = \begin{cases} \lambda_m & \text{if } a_t = y_t \\ -\lambda_k & \text{if } a_t \neq y_t \end{cases}$$

Where  $\lambda_m$  and  $\lambda_k$  are inversely proportional to the presence of the class and protected group in the training set:

$$\lambda_k = \frac{1}{N_k} \left( \frac{1}{N_0}, \frac{1}{N_1}, \dots, \frac{1}{N_k} \right)$$

This ensures that misclassifications for minority classes and protected groups receive higher penalties, thus encouraging the agent to learn fairer decision boundaries.

##### 5.3 Training Procedure

We employ a Double DQN approach to avoid the overestimation of Q-values, which can lead to suboptimal policy learning. During training, two separate networks are maintained: the target network and the main network. The target network is used for action-value estimation, while the main network updates the policy. The target:

$$y_i = r_t + \gamma Q(s_{t+1}, \arg \max_a Q(s_{t+1}, a; \theta); \theta^-)$$

Where  $\theta$  and  $\theta^-$  represent parameters of main and target networks, respectively. By decoupling action selection and evaluation, we ensure more stable learning and reduce bias in reward propagation.

## 6. EXPERIMENTAL SETUP

### 6.1 Implementation

We implemented the proposed deep reinforcement learning (RL) framework using Python and TensorFlow. The experiments were conducted on an NVIDIA Tesla V100 GPU with 32 GB memory, alongside a system with 256 GB of RAM and Intel Xeon processors. Our RL-based bias mitigation model was trained for hierarchical microcredential classification tasks using a dueling double-deep Q-network (Dueling DDQN) architecture. The model leveraged structured student data and narrative embeddings derived from hierarchical microcredentials. To ensure model convergence, we trained the architecture for approximately 4,000 steps. For optimization, we used Adam Optimizer with a LR of 0.0001, incorporating gradient clipping to prevent exploding gradients. We applied an  $\epsilon$ -greedy policy for exploration-exploitation, with  $\epsilon$  decaying linearly from 1.0 to 0.01 over 2,000 steps. Grid search with five fold CV selected the final hyper parameters e.g.  $\eta = 1 \times 10^{-4}$ , batch 64, hidden 512, dropout 0.2,  $\gamma = 0.99$  maximizing F1 while keeping equalized odds  $< 0.05$ . All experiments run reproducibly in a Docker image on an NVIDIA A100 (80 GB, host RAM 256 GB); make reproduce regenerates results in 10h. Training uses a 1000 step target network update,  $\Delta t = 0.01$  threshold grid, and early stopping after five stagnant epochs; the ADV baseline is trained in same environment for fair comparison (Table 4).

Table 4: Hyperparameters for proposed classifier and RL debiaser

Parameter	Search Range	Chosen Value
Learning rate $\eta$ (Adam)	{1e-5, 1e-4, 5e-4}	1e-4
Batch size	{32, <b>64</b> , 128}	64
Transformer hidden size	{256, 384, <b>512</b> }	512
Dropout $p$	{0.0, 0.1, <b>0.2</b> , 0.3}	0.2
RL discount $\gamma$	{0.90, 0.95, <b>0.99</b> }	0.99
Replay buffer size	{5M, <b>10M</b> }	10 M
Net update $C$ (steps)	{500, <b>1 000</b> , 2 000}	1 000
$\epsilon$ greedy start $\epsilon_0$	<b>1.0</b> (fixed)	1.0
$\epsilon$ decay (steps)	{1 000, <b>2 000</b> , 4 000}	2 000
Weight decay $\lambda$	{0, 1e-5, <b>5e-5</b> }	5e-5
Random seed	42 (all runs)	

**Baselines:** We benchmark our method against (i) a vanilla RoBERTa classifier and (ii) an adversarial debiasing model et al. [36]. Reinforcement learning is selected over re weighting or post processing approaches because its reward function lets us jointly optimize accuracy and Equalized Odds during training, yielding adaptive bias-utility trade offs without retraining for each fairness target.

## 7. RESULTS

### 7.1 Dynamic Thresholding Based Hierarchical Classification Results

In this section, we evaluate the effectiveness of our proposed RoBERTa model with dynamic thresholding for hierarchical classification of microcredentials. We compare it against the baseline RoBERTa model without dynamic thresholding. The evaluation is performed across all three levels of hierarchical microcredential classification (Level 1, Level 2, and Level 3), using standard classification metrics: Sensitivity, Specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), F1 Score, and Area Under the Receiver Operating Characteristic Curve (AUROC).

RoBERTa with dynamic thresholding outperforms the baseline model

across all hierarchical levels in terms of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), F1 score, and area under the receiver operating characteristic curve (AUROC). At Level 1, the dynamic thresholding model achieves a sensitivity of 0.892 and a specificity of 0.725, compared to the baseline’s 0.835 and 0.675, respectively. This trend continues at Level 2, with improvements in sensitivity (0.885 vs. 0.823) and specificity (0.710 vs. 0.658), and at Level 3, where dynamic thresholding yields a sensitivity of 0.878 and specificity of 0.700, surpassing the baseline’s 0.810 and 0.645. PPV is higher with dynamic thresholding at all levels (0.295, 0.285, and 0.275) compared to the baseline (0.245, 0.235, and 0.225), while NPV also improves (0.980, 0.975, and 0.970 vs. 0.965, 0.960, and 0.955). The F1 score benefits similarly, with values of 0.420, 0.410, and 0.400, compared to the baseline’s 0.375, 0.360, and 0.350. Lastly, the dynamic thresholding model achieves higher AUROC scores across all levels (0.830, 0.825, and 0.815), outperforming the baseline’s 0.795, 0.785, and 0.775, demonstrating superior classification performance and robustness in handling hierarchical microcredential tasks. The results in Table 5 clearly demonstrate that RoBERTa with dynamic thresholding consistently outperforms the baseline RoBERTa model across all evaluation metrics and all hierarchical levels. Dynamic thresholding allows for more fine-grained control over classification thresholds at each level, resulting in improved sensitivity, specificity, PPV, NPV, F1, and AUROC. These improvements highlight the effectiveness of the dynamic thresholding method in addressing the challenges of hierarchical multi-label classification in complex real-world datasets, further establishing its utility for tasks such as microcredential classification.

### 7.2 Debiasing Results

We introduced a method for training fair, unbiased machine learning (ML) models using a deep reinforcement learning (RL) framework. Our evaluation focused on hierarchical microcredential classification tasks, specifically targeting the reduction of gender (with males as the privileged group) and ethnicity/race biases (with white as the privileged group). We compared the proposed RL method with an adversarial debiasing (ADV) method, analyzing their performance across various metrics, including sensitivity, specificity, equalized odds (EO) for true positive (TP) and false positive rates, positive predictive value (PPV), negative predictive value (NPV), F1 score, and area under the receiver operating characteristic curve.

**Debiasing Gender:** We evaluated gender bias with males as the privileged group, and as shown in Table 5, both RL and ADV models demonstrated strong classification performance with high AUROC scores across all levels of hierarchical microcredentials. For Level 1, RL achieved the best equalized odds (EO) with EO(TP) = 0.030 and EO(FP) = 0.025, while ADV scored EO(TP) = 0.041 and EO(FP) = 0.029. RL also achieved a sensitivity of 0.892, with a slightly lower specificity (0.560) than ADV (specificity 0.630), which had a lower sensitivity (0.881). At Level 2, RL continued to outperform ADV in EO, with EO(TP) = 0.031 and EO(FP) = 0.026, compared to ADV’s EO(TP) = 0.040 and EO(FP) = 0.028, with both models performing similarly in sensitivity and ADV maintaining slightly higher specificity. At Level 3, RL maintained its lead in EO (EO(TP) = 0.029, EO(FP) = 0.027) over ADV (EO(TP) = 0.042, EO(FP) = 0.030). Overall, RL consistently achieved better EO across all levels, indicating superior fairness in mitigating gender bias while maintaining robust classification performance.

**Debiasing Ethnicity/Race:** In evaluating ethnicity/race bias, with white as the privileged group, RL consistently outperformed the ADV model in equalized odds (EO) across all classification levels,

Table 5: Comparison of RoBERTa without and with Dynamic Thresholding for Hierarchical Classification of Microcredentials

Model	Sensitivity	Specificity	PPV	NPV	F1	AUROC
<b>Level 1 Microcredentials Classification</b>						
RoBERTa without Dynamic Thresholding	0.835 ( $\pm 0.007$ )	0.675 ( $\pm 0.006$ )	0.245 ( $\pm 0.008$ )	0.965 ( $\pm 0.004$ )	0.375	0.795 ( $\pm 0.009$ )
RoBERTa with Dynamic Thresholding	0.892 ( $\pm 0.006$ )	0.725 ( $\pm 0.007$ )	0.295 ( $\pm 0.007$ )	0.980 ( $\pm 0.003$ )	0.420	0.830 ( $\pm 0.008$ )
<b>Level 2 Microcredentials Classification</b>						
RoBERTa without Dynamic Thresholding	0.823 ( $\pm 0.008$ )	0.658 ( $\pm 0.007$ )	0.235 ( $\pm 0.009$ )	0.960 ( $\pm 0.005$ )	0.360	0.785 ( $\pm 0.009$ )
RoBERTa with Dynamic Thresholding	0.885 ( $\pm 0.007$ )	0.710 ( $\pm 0.008$ )	0.285 ( $\pm 0.008$ )	0.975 ( $\pm 0.004$ )	0.410	0.825 ( $\pm 0.009$ )
<b>Level 3 Microcredentials Classification</b>						
RoBERTa without Dynamic Thresholding	0.810 ( $\pm 0.008$ )	0.645 ( $\pm 0.007$ )	0.225 ( $\pm 0.009$ )	0.955 ( $\pm 0.005$ )	0.350	0.775 ( $\pm 0.010$ )
RoBERTa with Dynamic Thresholding	0.878 ( $\pm 0.007$ )	0.700 ( $\pm 0.008$ )	0.275 ( $\pm 0.009$ )	0.970 ( $\pm 0.004$ )	0.400	0.815 ( $\pm 0.009$ )

Table 6: Microcredentials Classification Results for Gender and Ethnicity/Race Bias (RL and ADV Models)

Model	EO (TP)	EO (FP)	Sensitivity	Specificity	PPV	NPV	F1	AUROC
<b>Level 1 Microcredentials Classification (Gender Bias, Male as Privileged Group)</b>								
RL	0.030a	0.025a	0.892 ( $\pm 0.005$ )	0.560 ( $\pm 0.006$ )	0.180 ( $\pm 0.004$ )	0.978 ( $\pm 0.003$ )	0.295	0.820 ( $\pm 0.008$ )
ADV	0.041	0.029	0.881 ( $\pm 0.006$ )	0.630 ( $\pm 0.007$ )	0.210 ( $\pm 0.009$ )	0.979 ( $\pm 0.003$ )	0.325	0.860 ( $\pm 0.006$ )
<b>Level 2 Microcredentials Classification (Gender Bias, Male as Privileged Group)</b>								
RL	0.031a	0.026a	0.890 ( $\pm 0.006$ )	0.545 ( $\pm 0.007$ )	0.185 ( $\pm 0.005$ )	0.979 ( $\pm 0.003$ )	0.290	0.822 ( $\pm 0.009$ )
ADV	0.040	0.028	0.880 ( $\pm 0.007$ )	0.635 ( $\pm 0.008$ )	0.215 ( $\pm 0.009$ )	0.980 ( $\pm 0.003$ )	0.320	0.858 ( $\pm 0.007$ )
<b>Level 3 Microcredentials Classification (Gender Bias, Male as Privileged Group)</b>								
RL	0.029a	0.027a	0.885 ( $\pm 0.007$ )	0.540 ( $\pm 0.006$ )	0.175 ( $\pm 0.006$ )	0.977 ( $\pm 0.004$ )	0.285	0.818 ( $\pm 0.008$ )
ADV	0.042	0.030	0.878 ( $\pm 0.007$ )	0.630 ( $\pm 0.007$ )	0.210 ( $\pm 0.008$ )	0.978 ( $\pm 0.003$ )	0.320	0.855 ( $\pm 0.007$ )
<b>Level 1 Microcredentials Classification (Ethnicity/Race Bias, White as Privileged Group)</b>								
RL	0.033a	0.024a	0.891 ( $\pm 0.005$ )	0.550 ( $\pm 0.006$ )	0.175 ( $\pm 0.004$ )	0.979 ( $\pm 0.003$ )	0.290	0.821 ( $\pm 0.008$ )
ADV	0.040	0.030	0.880 ( $\pm 0.006$ )	0.620 ( $\pm 0.007$ )	0.200 ( $\pm 0.009$ )	0.978 ( $\pm 0.003$ )	0.320	0.859 ( $\pm 0.006$ )
<b>Level 2 Microcredentials Classification (Ethnicity/Race Bias, White as Privileged Group)</b>								
RL	0.034a	0.027a	0.889 ( $\pm 0.006$ )	0.540 ( $\pm 0.007$ )	0.180 ( $\pm 0.005$ )	0.978 ( $\pm 0.003$ )	0.285	0.823 ( $\pm 0.009$ )
ADV	0.038	0.029	0.878 ( $\pm 0.007$ )	0.630 ( $\pm 0.008$ )	0.210 ( $\pm 0.009$ )	0.980 ( $\pm 0.003$ )	0.315	0.857 ( $\pm 0.007$ )
<b>Level 3 Microcredentials Classification (Ethnicity/Race Bias, White as Privileged Group)</b>								
RL	0.032a	0.026a	0.884 ( $\pm 0.007$ )	0.535 ( $\pm 0.006$ )	0.170 ( $\pm 0.006$ )	0.976 ( $\pm 0.004$ )	0.280	0.817 ( $\pm 0.008$ )
ADV	0.041	0.031	0.877 ( $\pm 0.007$ )	0.620 ( $\pm 0.007$ )	0.205 ( $\pm 0.008$ )	0.977 ( $\pm 0.003$ )	0.315	0.854 ( $\pm 0.007$ )

as shown in Table 6. For Level 1, RL achieved superior EO(TP) = 0.033 and EO(FP) = 0.024, compared to ADV's EO(TP) = 0.040 and EO(FP) = 0.030, with both models maintaining high sensitivity (RL: 0.891, ADV: 0.880) and RL having slightly lower specificity (0.550 vs. ADV's 0.620). At Level 2, RL again led in EO scores (EO(TP) = 0.034, EO(FP) = 0.027) compared to ADV (EO(TP) = 0.038, EO(FP) = 0.029), with similar sensitivity for both models but slightly lower specificity for RL. For Level 3, RL demonstrated the best EO scores (EO(TP) = 0.032, EO(FP) = 0.026), surpassing ADV (EO(TP) = 0.041, EO(FP) = 0.031), and maintained high sensitivity (RL: 0.884, ADV: 0.877) despite a marginal trade-off in specificity. Overall, RL consistently demonstrated better fairness in mitigating ethnicity/race bias across all classification levels.

### 7.3 Model Generalization and Trade-offs

Both RL and ADV models exhibited strong classification performance, but RL consistently demonstrated better fairness, as indicated by improved equalized odds in true positive and false positive rates. This advantage was clear in both gender and ethnicity/race bias mitigation tasks. While both models achieved high AUROC scores across hierarchical microcredential classification levels, RL showed slightly lower specificity compared to ADV. However, this trade-off was compensated by RL's superior fairness and sensitivity scores. Overall, RL proved robust and consistent, maintaining comparable sensitivity, specificity, and AUROC values to ADV while significantly improving equalized odds. This highlights RL's effectiveness in achieving fairer outcomes without notable performance degradation, making it an effective solution for mitigating bias in complex classification tasks. The difference in fairness (measured by equalized odds) between the RL and ADV models was statis-

tically significant across both gender and ethnicity/race bias mitigation tasks ( $P < 0.001$ , by the Wilcoxon signed rank test). RL consistently outperformed ADV in equalized odds, particularly in terms of EO for true positive and false positive rates across all levels of classification. This statistical significance underscores the robustness of the RL model in mitigating bias.

## 8. CONCLUSION

We introduced bias free intelligent deep learning framework for hierarchical microcredential classification, combining dynamic thresholding with reinforcement learning to mitigate gender and ethnicity/race biases. By incorporating a dueling double-deep Q-network (Dueling DDQN), our method addresses algorithmic biases that are prevalent in hierarchical multi-label classification tasks. Although Dueling DDQN is a known RL architecture, this work is the first to adapt it for hierarchical micro credential NLP and to embed an equalized odds fairness reward, making the algorithm bias aware rather than purely performance driven. Through extensive experiments on student lived experience narratives, we demonstrated that the proposed framework significantly outperforms traditional approaches in both classification accuracy and fairness across all levels of microcredential hierarchy. Despite the promising results, this work faces limitations, including potential scalability challenges when applied to larger datasets or more complex hierarchical structures. While effective on a 3,000-student narrative dataset, scaling up may require significant computational resources and optimization. Additionally, reliance on human-annotated data introduces a bottleneck, as annotations are time-consuming and can still carry subjective biases. Overall, the paper offers a strong foundation for the development of fair, scalable, and accurate microcredential classification systems.

## References

- [1] M. A. U. Alam, M. Pagare, S. Davis, G. Verma, A. K. Biswas, and J. Barbern. Empowering predictions of the social determinants of mental health through large language model augmentation in students' lived experiential essays. In *Proceedings of the 17th International Conference on Educational Data Mining, EDM 2024, Atlanta, Georgia, USA, July 14-17, 2024*. International Educational Data Mining Society, 2024.
- [2] R. Binns. Fairness in machine learning: Lessons from political philosophy. 81:149–159, 23–24 Feb 2018.
- [3] A. K. Biswas, G. Verma, and J. O. Barber. Improving ethical outcomes with machine-in-the-loop: Broadening human understanding of data annotations. *CoRR*, abs/2112.09738, 2021.
- [4] T. Bolukbasi and K.-W. Chang. Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in Neural Information Processing Systems (NeurIPS)*, 29(1):4356–4364, 2016.
- [5] M. Brown, M. Nic Giolla Mhichil, E. Beirne, and C. Mac Lochlainn. The global micro-credential landscape: Charting a new credential ecology for lifelong learning. *Journal of Learning for Development*, 8(2):228–254, 2021.
- [6] Center for the Future of Higher Education and Talent Strategy. Getting started with micro-credentials: A primer for higher education leaders. Technical report, Northeastern University, 2022.
- [7] R. Cheng, H. Ma, S. Cao, and T. Shi. RLRf: reinforcement learning from reflection through debates as feedback for bias mitigation in llms. *CoRR*, abs/2404.10160, 2024.
- [8] S. J. Correll. Reducing gender biases in modern workplaces: Organizational approaches and their limitations. *Gender & Society*, 31(6):725–750, 2017.
- [9] S. DeMark. Skills-based education for the future of work. *Workforce Intelligence & Credential Integrity*. Accessed May 10, 2025.
- [10] Digital Promise. Can ai be leveraged to support competency-based assessments? 2023.
- [11] EDUCAUSE. Microcredentialing. 2025. Accessed May 10, 2025.
- [12] European Commission. Towards a european approach to micro-credentials. Technical report, Directorate-General for Education, Youth, Sport and Culture, 2019. Analytical Report.
- [13] European MOOC Consortium. Common micro-credential framework (cmf). <https://www.futurelearn.com/info/the-common-microcredential-framework>, 2019.
- [14] European Training Foundation. Guide to design, issue and recognise micro-credentials. <https://www.etf.europa.eu/sites/default/files/2023-05/Micro-Credential%20Guidelines%20Final%20Delivery.pdf>, 2023.
- [15] FutureLearn. Applications of machine learning - micro-credential. <https://www.futurelearn.com/microcredentials/applications-of-machine-learning>, 2024.
- [16] P. Gajane, A. Saxena, M. Tavakol, G. Fletcher, and M. Pechenizkiy. Survey on fair reinforcement learning: Theory and practice. *ACM Computing Surveys*, 2022.
- [17] M. Heidegger. *Being and Time*. Blackwell, Oxford, 1962.
- [18] L. Henrickson, D. O'Neill, and J. McCarthy. Soft skills, stories, and self-reflection: Applied digital storytelling for enhancing young adults' soft skills. *Convergence: The International Journal of Research into New Media Technologies*, 28(6):1575–1593, 2022.
- [19] McGraw-Hill Education. 2018 future workforce survey, 2018. Available at: <https://www.mheducation.com/unitas/corporate/promotions/2018-future-workforce-survey-analysis.pdf>.
- [20] N. Mehrabi et al. A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6):1–35, 2021.
- [21] L. C. Moll, C. Amanti, D. Neff, and N. Gonzalez. Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into Practice*, 31(2):132–141, 1992.
- [22] Nova Academy. Micro-credential machine learning of natural language processing. <https://nova-academy.be/en/programmes/machine-learning-of-natural-language-processing>, 2025.
- [23] OECD. Micro-credential innovations in higher education: Who, what and why? *OECD Education Policy Perspectives*, 2021.
- [24] S. Parisot and et al. Algorithmic fairness and bias mitigation for clinical machine learning: a reinforcement learning approach. *Nature Machine Intelligence*, 5:686–698, 2023.
- [25] C. Paulson, S. Panke, P. Carlson, and P. Roesler. Designing and implementing a micro-credentialing system for a local government leadership training program. pages 177–186, July 2023. University of North Carolina at Chapel Hill.
- [26] N. A. Saputra, I. Hamidah, and A. Setiawan. A bibliometric analysis of deep learning for education research. *Journal of Engineering Science and Technology*, 18(2):1258–1276, 2023.
- [27] R. M. Selvaratnam and M. D. Sankey. An integrative literature review of the implementation of microcredentials in higher education: Implications for practice in australasia. *Journal of Teaching and Learning for Graduate Employability*, 12(1):1–17, 2021. Corresponding author: Ratna Selvaratnam (r.selvaratnam@ecu.edu.au).
- [28] SUNY. Data mining and machine learning - microcredentials. <https://www.suny.edu/microcredentials/programs-test/1000000523/>, 2004.
- [29] N. Swaminathan and D. Danks. Bias mitigation via compensation: A reinforcement learning perspective. *CoRR*, 2024.
- [30] UNESCO Institute for Lifelong Learning. Short courses, micro-credentials and flexible learning pathways. <https://unesdoc.unesco.org/ark:/48223/pf0000384326>, 2023.



- [31] S. D. Walter, M. Eliasziw, and A. Donner. Sample size and optimal designs for reliability studies. *Statistics in Medicine*, 17(1):101–110, 1998.
- [32] Y. Wang, W. Gan, J. Yang, W. Wu, and J. Yan. Dynamic curriculum learning for imbalanced data classification. pages 5016–5025, 2019.
- [33] Z. Wang and et al. Reinforcement learning from multi-role debates as feedback for bias mitigation in llms. *arXiv preprint arXiv:2404.10160*, 2024.
- [34] Z. Wang, T. Schaul, M. Hessel, H. Van Hasselt, M. Lanctot, and N. De Freitas. Dueling network architectures for deep reinforcement learning. page 1995–2003, 2016.
- [35] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy. Hierarchical attention networks for document classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1480–1489, 2016.
- [36] B. H. Zhang, B. Lemoine, and M. Mitchell. Mitigating unwanted biases with adversarial learning. pages 335–340, 2018.