AI-Moderated Decision-Making: Capturing and Balancing Anchoring Bias in Sequential Decision Tasks

Jessica Maria Echterhoff
jechterh@eng.ucsd.edu
University of California, San Diego
USA

Matin Yarmand
myarmand@eng.ucsd.edu
University of California, San Diego
USA

Julian McAuley
jmcauley@eng.ucsd.edu
University of California, San Diego
USA

ABSTRACT

Decision-making involves biases from past experiences, which are difficult to perceive and eliminate. We investigate a specific type of anchoring bias, in which decision-makers are anchored by their own recent decisions, e.g. a college admission officer sequentially reviewing students. We propose an algorithm that identifies existing anchored decisions, reduces sequential dependencies to previous decisions, and mitigates decision inaccuracies post-hoc with 2% increased agreement to ground-truth on a large-scale college admission decision data set. A crowd-sourced study validates this algorithm on product preferences (5% increased agreement). To avoid biased decisions ex-ante, we propose a procedure that presents instances in an order that reduces anchoring bias in real-time. Tested in another crowd-sourced study, it reduces bias and increases agreement to ground-truth by 7%. Our work reinforces individuals with similar characteristics to be treated similarly, independent of when they were reviewed in the decision-making process.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; • **Social and professional topics** → *Computing / technology policy*; • **Applied computing** → Law, social and behavioral sciences.

KEYWORDS

Anchoring, Bias, Neural Networks, Human-AI Interaction, Decision-Making

ACM Reference Format:
Jessica Maria Echterhoff, Matin Yarmand, and Julian McAuley. 2022. AI-Moderated Decision-Making: Capturing and Balancing Anchoring Bias in Sequential Decision Tasks. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3491102.3517443

1 INTRODUCTION

Consider the following scenario: A reviewer evaluates many unqualified applicants for a university program successively, and the next applicant to be reviewed is an average (borderline admit-reject) applicant. Because the evaluator is influenced, or **anchored**, by recently made decisions, this borderline applicant might be admitted to the program. On the other hand, when an evaluator is anchored by having reviewed many qualified applicants, the same or a similarly borderline applicant might be rejected (Figure 1). In this scenario, *individual fairness*, stating that individuals with similar characteristics should be treated similarly [8], is impaired, and wrong or inconsistent decisions can have a consequential impact.

Many decision-making tasks are subject to this form of anchoring bias. Human evaluators are often required to sequentially review instances as subject-matter experts, for example, to select submitted papers for publication to a conference, select job applicants [19, 35, 38], or to make bail or sentencing decisions [21]. Those anchoring effects are observed in both high-stakes decision tasks such as in the evaluation of graduate student applications for a graduate program of a large US university, as well as in low-stakes decision tasks, such as the sequential evaluation of products by crowd workers.

1.1 Motivation

Our work was motivated by the observation that there is a correlation between 1) the number of decisions made by a human evaluator since the last positive evaluation, and 2) the confidence of a process that makes decisions independently of the position in a review sequence. This could be e.g. a Machine Learning (ML) algorithm like a Support Vector Machine (SVM). When the SVM learns from the data, decision boundaries are formed to identify which class a sample belongs to. The SVM confidence describes the distance to this decision boundary.

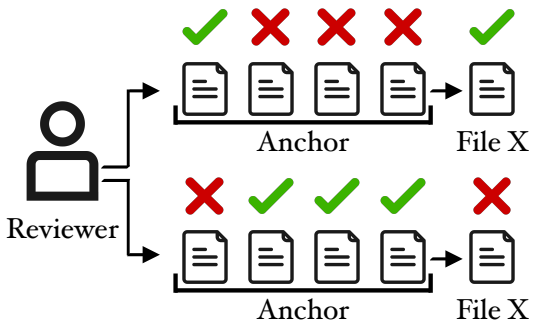


Figure 1: In sequential decision making tasks, previously made decisions can influence, or *anchor*, the next decision of a human reviewer. Those anchors can lead to incorrect or inconsistent decisions of the same reviewed instance X.

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04.
https://doi.org/10.1145/3491102.3517443

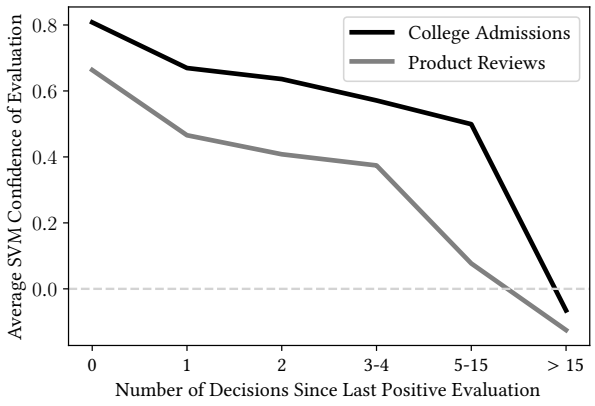


Figure 2: In sequential decision-making tasks, human decision-makers appear to be influenced by their previously made decisions. This is indicated by decreasing confidence of an SVM, which is not subject to anchoring bias, the more negative decisions a human decision-maker makes in a row. Anchored human decisions lead to potentially wrong acceptance of low-quality instances.

The larger the distance, the more clear it is that this sample belongs to a specific class. Figure 2 shows a loss in SVM confidence that decreases the longer the last positive review is in the past. This behavior implies tendency of anchoring. The longer the last positive review of a human evaluator lies in the past, the more ambiguous the human decision is to the SVM, ultimately leading to the SVM disagreeing with the human decision. When the confidence is negative, the SVM predicts a different class than the human did.

This observation led to the hypothesis that past decisions influence the current judgment in such sequential decision-making tasks. If the last positive decision is too far in the past, it indicates that humans are naturally subject to be anchored by their previously made decisions and might tend to accept low-quality instances. Our work proposes and evaluates procedures to mitigate this specific form of anchoring bias to increase fairness in decision-making.

1.2 Contributions

Our work learns and mitigates anchoring bias, such that enhanced knowledge about systematic biases in decision-making can be leveraged to support people in making more optimal decisions.

1.2.1 Observing anchoring effects in sequential decision-making. We study data from decisions made within a college admissions process to a graduate program of a large US university, approved to be studied by our university. We show findings of anchoring bias in sequentially made admission decisions. In a crowdsourced experiment on the sequential evaluation of product reviews, we find similar anchoring tendencies. Unlike prior papers that focus on anchoring bias induced by task features [10, 15, 28, 34–36], this work provides results on anchoring bias induced by previously made decisions.

1.2.2 Moderating sequences to mitigate anchoring effects. We propose two bias mitigation strategies for sequential decision-making tasks when human evaluators appear to be influenced by their previous decisions. 1) To *retrospectively mitigate anchoring bias* when the decision-making process is completed, and all instances were sequentially reviewed, we capture a user's *anchoring state* within a probabilistic model. When an evaluator appears to be anchored, we adapt the decision that was made and observe the change in accuracy with respect to the underlying ground-truth decisions. Our procedure helps to detect biased decisions when there is no influence over the decision-making process, or influencing the decision process directly could otherwise raise ethical concerns. Our proposed algorithm increases accuracy to ground-truth decisions by 2-5% and reduces bias by 0.01-0.08, measured with the Pearson correlation coefficient. 2) To *prospectively mitigate anchoring bias*, when a reviewer is in the process of sequentially evaluating instances, we propose an algorithm to learn how to choose the next instance to be shown to an evaluator. The algorithm learns the current *anchoring state* with a parametric model and uses this state to learn which instance to show next in order to minimize anchoring effects. This procedure is evaluated in a crowdsourced experiment on product reviews. The procedure reduces sequential dependencies on previous decisions by 0.07 (Pearson correlation coefficient) and increases agreement to ground-truth by 7%. This algorithm helps to prevent anchored decisions before the decisions are made, and can be used to mitigate bias within a live, interactive process to increase fairness in decision-making.

2 RELATED WORK

2.1 Anchoring Bias is Present in Human Decision Tasks

If you first see a T-shirt that costs $1,000, and then see a second one that costs $100, you're prone to see the second shirt as cheap. This behavior is called *anchoring bias* [36] and studied for a variety of application areas like recommender systems or Human-Artificial Intelligence (AI) decision-making. For a decision to be made, both the setup of the task and information shown within the task can be an anchor for the decision-maker. For example, in recommender system research, showing ratings of an object to a user significantly influences user preferences for products [1]. In marketing, displaying an old, higher price as the anchor gives the impression of a better deal to an evaluator [3]. In Human-AI decision-making, the decision of the algorithm shown to a user as additional information [2] can anchor the evaluator. Anchoring bias can influence the consistency of evaluations, meaning a user's answer differs depending on if they were anchored or not, e.g. when evaluating conversational agents with the crowd [32]. To mitigate this kind of bias, researchers tried to make users aware of their own bias e.g. with a note or warning screen within the decision-making process [4, 11]. This procedure requires an adaptation of the task and can itself impose a risk of anchoring the evaluator. Our study is distinct in its approach to mitigating anchoring bias without showing any additional information to a user.

2.2 Anchors can be Induced by Previous Decisions

In comparison to anchors induced by the task setup, evaluators making repeated sequential decisions can be anchored by their own recent decisions [23, 39]. Evidence of bias in sequential decision-making has been found for subjective judgments like rating face attractiveness [12] loudness [14] or within the review process for undergraduate admission decisions [35]. Mitigating those biases can be seen as *nudging*, where subtle changes in the 'choice architecture' can alter people's behaviors in predictable ways [6, 18]. Those nudges are frequently used in decision-making [7, 16, 20, 33, 41, 43]. For example, *Talkad et al.* [35] presented a mitigation strategy that re-arranges the order of attributes of a single student instance that is presented to an evaluator. This enables the non-sequential consideration of conjunctive attributes that are affected by the anchor. This procedure requires knowledge on which instance attributes are affected by the anchor, which is often non-trivial to obtain. *Huang et al.* [12] mitigate bias in sequential decision tasks by alternating stimuli, like speech and images, when different modalities are available for decision-making of attractiveness. This process requires understanding the relationship of anchoring to different modalities which might be highly specific to the decision task and impossible to use in a single-modality task setup. In comparison, our method requires a setup where we can train a classifier, which can often be learned for a variety of tasks. Our procedure does not require other domain- or modality-specific knowledge. Compared to previous work, our methods *nudge* a reviewer towards less biased decisions, without interfering in the decision process or adapting the decision setup [4, 11, 12, 35].

2.3 Individuals might be Anchored Differently

Studies on anchoring typically observe bias within a given task and then adapt the procedure without accommodating differences between the individual subjects [35, 36]. To work towards more personalized mitigation of anchoring bias, *George et al.* [11] heuristically estimated the anchoring state of an evaluator. They then guide the solution to the task by providing domain knowledge to the user depending on whether anchoring was detected. We take this approach further by specifically learning the anchoring state of an evaluator with probabilistic and ML models instead of heuristically estimating it.

3 MITIGATING ANCHORING BIAS IN SEQUENTIAL DECISION TASKS

To mitigate anchoring bias from *highly accurate but potentially biased human decisions*, we require a *method not subject to anchoring bias*. Different ML algorithms can be used to serve as such a method. These do not have access to the specific ordering of the files and hence are not subject to sequential dependencies that could cause anchoring. We decide to train a Support Vector Machine (SVM) predictor for this purpose. For college admissions, the data set is large enough to learn admission predictions from the 305 student features (e.g. GPA scores) (accuracy 98%). For product reviews, the SVM predicts if a product is reviewed positively or negatively with an accuracy of 77% on the book review test set [24]. TF-IDF word encodings [27] were used to featurize the review sentences, which use a Bag-of-Words approach to get a sparse, fixed-length vector representation of the sentences for further SVM training. The product review SVM is trained using additional review data from music, Amazon Kindle, and book reviews [24], which leads to a performance gain compared to training only on book reviews. In total, over 69,000 review sentences were used to train the product review SVM. The SVM decision and its confidence, measured in the distance to the decision boundary, are further used to aid the mitigation process.

3.1 Retrospectively Mitigating Anchoring Bias with Probabilistic Adaptation (PA)

We assume an ML algorithm, like an SVM, to not be influenced by anchoring bias, because it cannot use the ordering of decisions made. Figure 1 shows a tendency that the decision outcome of an evaluator is influenced by the previously made decisions, where we see an exponential decay of SVM confidence with respect to the number of decisions that were made since the last acceptance. This indicates a correlation between the anchoring state of a reviewer and the number of decisions made since the last positive evaluation. Based on this observation, we fit a probabilistic exponential decay function to the number of decisions that were made since the last acceptance. This function f is fit to our data using non-linear least squares, where x indicates the number of decisions made since the last acceptance and λ is the parameter that is fitted to our actual data (Equation 1).

$$f(x;\lambda) = \lambda e^{-\lambda x} \tag{1}$$

Based on the qualitative review on the shape of the SVM confidence, we decided on the general shape of the model to be approximated by the probabilistic exponential decay function. The final function reflects a probability that the reviewer is biased, given the number of decisions that were made since the last acceptance, and serves as an indication of the anchoring state of the evaluator. We draw a sample from a binomial distribution with probability $p_{anchored} = 1 - f(x;\lambda)$, as the probability of a human being anchored. The decision is adjusted using a logical strategy depending on the human decision d_{human}, the outcome of the drawing $d_{bias} \sim Bin(p_{anchored})$, and the decision of the SVM estimator d_{SVM}. The decision is then adjusted with the help of the SVM when an evaluator appears to be biased:

$$d_{adjusted} = (d_{bias} \wedge d_{SVM}) \vee d_{human}. \tag{2}$$

Equation 2 is used to explain how the decision of a reviewer can be adjusted retrospectively depending on the decision of the SVM, the decision of the human reviewer, and the probability of the human reviewer being anchored by their previous decisions.

3.2 Learning a Review Sequence Order that Mitigates Anchored Decisions

We aim to learn in which order instances should be reviewed to minimize anchoring bias for evaluators. To do this, we explore different strategies. First, we aim to understand whether anchoring bias in sequential decision tasks can be mitigated using heuristic strategies. If heuristics are sufficient to mitigate anchoring bias, no complicated ML methods would be required. We experiment with a heuristic strategy that alternately shows strong and weak instances

to evaluators, depending on their previously made decision. Second, we use ML to learn which instances to show next to an evaluator. For this procedure, we need a component to simulate an evaluator's decisions and indicate the anchoring state of a reviewer. We subsequently need a procedure to learn how to decide on the next instance to be shown. Those components are shown in Figure 3 and explained in the next sections.

3.3 Capturing Anchoring Bias with Long-Short-Term-Memory (LSTM) Neural Networks

To mitigate human anchoring bias before decisions are made, a procedure is required to surface instances to an evaluator in an order that minimizes potential anchoring. This requires an algorithm to choose instances in an order that will cause users to make decisions independent of their previously made decisions. To decide on the next instance to be shown, we need to quantify if an evaluator is anchored or not, and the degree of the anchoring state. We use an LSTM neural network to get a **measure of the anchoring state** of an evaluator in a particular review sequence. This network simulates human decisions for the review sequence and captures an evaluator's anchoring state within its hidden state. Since there might be differences in the amount of anchoring experienced by a human evaluator, we find it to be of importance to model the actual decisions made, rather than only modeling the number of decisions made since the last positive evaluation. That way, we get the extent of anchoring of an evaluator within the LSTM's hidden state. The previously made decisions by the evaluator $d_{t_0,\ldots,t_{i-1}}$ for time steps $t_0, \ldots, t_{i-1}$ as well as the predictions from the SVM are fed into the LSTM for all previously made decisions of the sequence. The network then produces the anchoring states $h_{t_0,\ldots,t_{i-1}}$ and simulated decisions $\hat{d}_{t_0,\ldots,t_i}$ using the LSTM output. The simulated decision output is a probability, obtained a using a linear layer and softmax in the network. Cross-entropy loss is used to update our model weights for the LSTM. The raw hidden states of the LSTM are used as an indicator of the state of anchoring for a review sequence. This component is important for simulating decisions and quantifying the anchoring state as shown in Figure 3, to be able to decide subsequently which instance to show next.

3.4 Mitigating Anchoring Bias with Parametric Reinforcement Learning (RL) Models

After quantifying the anchoring state of the evaluator, instances have to be shown to a user in an order which promotes un-anchored decisions. The order which minimizes anchoring effects is not easily quantifiable, and hence cannot be learned by supervised learning procedures. We use the anchoring state produced by the LSTM network as the input for the bias mitigation procedure with RL algorithms. RL methods allow us to define the goal of *minimizing the overall anchoring state.* With this goal, we can learn how to decide on the *order of the instances to be shown to a reviewer* given a particular anchoring state within a review session. Given the hidden anchoring states $h_{t_0,\ldots,t_{i-1}}$ for a review sequence predicted by the LSTM network, we use two RL algorithms for anchor mitigation (Figure 3). These Actor-Critic (AC) [30] and Deep-Q-Network (DQN) [22] RL algorithms decide on the next instance to be surfaced

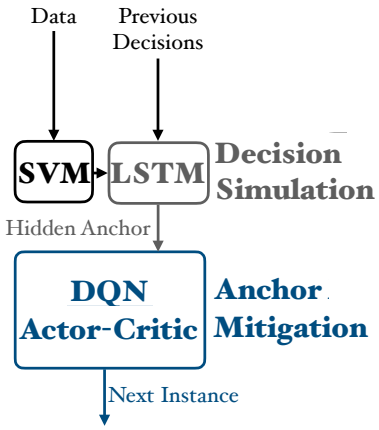


Figure 3: Architecture overview of the system mitigating anchoring bias. We learn if an evaluator is biased by making decision predictions with an SVM, passing them to an LSTM neural network, and generating an anchoring state of the evaluator. We decide on the best instance to show next to an evaluator, by passing the anchoring state to an anchor mitigation (RL) algorithm which decides on an instance to show next, that minimizes the overall anchoring in a review sequence.

to minimize anchoring, based on the anchoring states $h_{t_0,\ldots,t_{i-1}}$ obtained from the LSTM neural network. The RL procedure learns how to mitigate anchoring by sampling specific kinds of instances to re-balance the anchoring state of the evaluator. Further technical implementation details can be found in the appendix (Section 8).

4 STUDY DESIGN

To evaluate the methods in our study, we use data from two distinct decision-making tasks. In both tasks, evaluators were asked to sequentially make decisions by rating instances. We analyze and mitigate anchoring effects on both data sets.

4.1 Data Set 1: College Admissions

Our first data set contains decisions on college admissions from a graduate program at a large US university. The collection and usage of this data were approved to be studied by the university. This data represents a common sequential decision task for which biases can be observed. In this decision task, anonymous academic reviewers ($N = 117$) rate college students based on their application materials submitted. No process determines the ordering of student instances for reviewers. A reviewer could review files in any order they like, similarly to paper reviews for a conference. Evaluators can choose how many students they review in succession and rate students on a scale of 0-3 (0 worst–3 best). Over 5,000 students apply annually to the graduate program, and each evaluator reviews around 50 students on average. The decisions are binarized to binary 0 (rejection to the program) when rated < 2 and 1 (admission to the program) otherwise. After all individual decisions are made, a committee decides whether or not to admit the student to the graduate program. This decision serves as the ground-truth decision

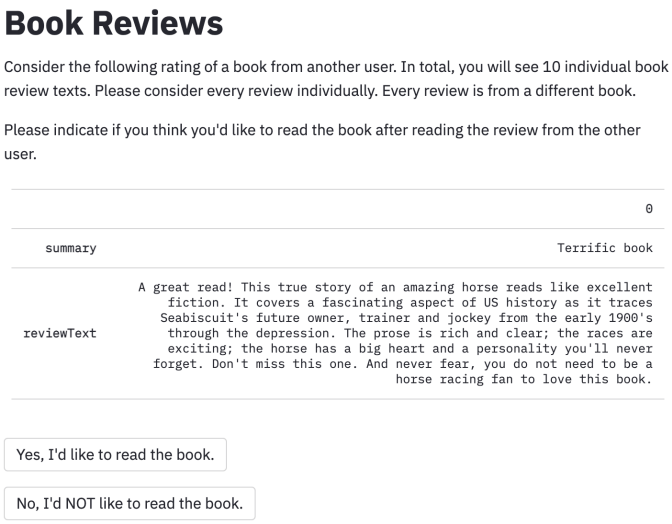


Figure 4: Interface of the user study presented to crowd workers in all of our experiments. Crowd workers were shown different product reviews sequentially. They had to decide if they were interested in the product or not. The order of the instances shown to participants varies depending on the experiment.

for our approaches. We assume this decision to be 'correct' for most of the evaluated students since the decision is analogous to a meta-review by a committee chair, who is not subject to the same anchor state. From this procedure, we extract the data set containing features of college applicants, like GPA or quality of the statement of purpose, the individual evaluator decision, timestamp of the decision, and the final committee decision.

4.2 Data Set 2: Product Reviews

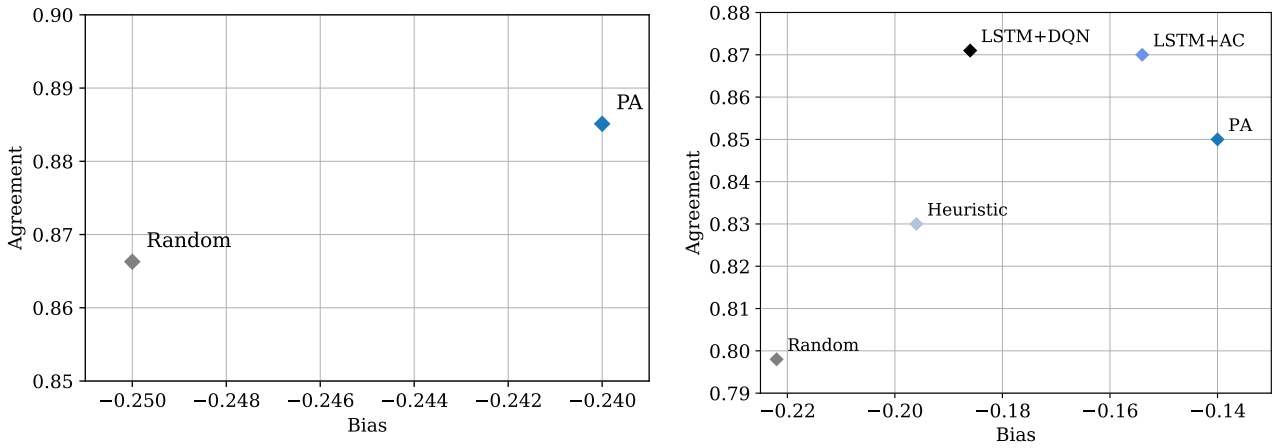
It is ethically questionable to run live experiments on a high-stakes decision task such as college admission decisions. To evaluate our live mitigation strategy, we designed an experiment with similar characteristics to the college admission decisions. In a typical admissions process, there are more students rejected than are admitted to a program, hence we chose our data set to have more negatively than positively attributed instances. We also specifically chose a review process that depends on the subjective preferences of an evaluator. For example, one college-admissions evaluator might prefer a strong GPA over a strong motivation letter; similarly, one evaluator in our product review setting might prefer one book genre over another. We collected human decisions on product preferences in a sequential decision-making task performed on Amazon Mechanical Turk. Crowd workers were asked to decide if they are interested in a product after showing an existing written review of this product from another user, drawn from a publicly available Amazon review data set [24]. This review data consists of a summary of the review, the review text, and a rating of the product on a scale from 1–5 (1 is the worst - 5 is the best rating) from the person who wrote the review. The evaluator was shown the summary and review text but was not exposed to the original scalar product rating, which serves as the ground-truth of if the review was positive or negative. We asked participants to decide on a binary scale whether or not they are interested in the product. We use the user interface shown in Figure 4 to get decisions from crowd workers (N=90). The files are shown to the user with this user interface for all experiments with the crowd. We deliberately choose a simplistic user interface to avoid priming the user as much as possible. To ensure the quality of the reviews and avoid confusion due to language barriers, only crowd workers with an approval rate of $> 95\%$ based in the United States or the United Kingdom were approved to participate in the study. On this data, we can run live experiments to test our bias mitigation strategies. In total, 436 unique product reviews were displayed to the study participants in different orders, depending on the mitigation strategy used. For the baseline review setup, the next instances to be shown in random order were pre-rendered before the study setup. Specifically, we used a random perturbation of the instances and a random length of the review sequence. The review sequences that were re-ordered by our mitigation strategies were shown to the crowd workers in the same setup shown in Figure 4, as they were for the baseline random order. Some, but not all of the crowd workers completed sequences from different algorithms.

4.3 Analysing Bias and Accuracy

To evaluate our methods it is necessary to measure the extent of anchoring bias that is present and accuracy of the decisions made. This requires quantifying anchoring bias with respect to the sequential decisions made by an evaluator. Since evaluators appear to be more biased the more decisions they have made since the last positive decision (Figure 2), we calculate the Pearson correlation coefficient between the number of decisions since the last positive decision was made and the current decision of the evaluator. This aims to be a quantifiable measure of bias in sequential decisions and is a common interpretable statistical metric that can be derived from the initial motivation of decreasing confidence of an ML method not subject to anchoring. To measure if our methods promoted less anchored decisions, we calculate the agreement and accuracy of human evaluators to ground-truth decisions. For the product review data, we use the scalar rating of the product from the person who wrote the product review as ground-truth. For the admissions data set, we use the final admission decision made by the committee as ground-truth. The agreement of the decision of an individual evaluator to this ground-truth is then computed.

5 RESULTS

Our collected college admission data set contains 26,174 decisions made within 5,814 review sequences by 117 unique evaluators. In our crowdsourced study on product reviews, we collected data from 145 review sequences, each consisting of 10–50 product reviews. In total, 3,570 decisions were made by 90 unique evaluators. As shown in the motivation for this work in Figure 2, human evaluators show a similar magnitude of anchoring bias with respect to their past decisions in both data sets. The data indicates that the number of decisions since the last positive decision has an impact on the outcome of the next decision. For both decision tasks, after more than 15 negative decisions, evaluators start to be more susceptible to misclassification.



(a) College Admissions **(b) Product Reviews**

Figure 5: Agreement to ground-truth and bias (Pearson correlation to previous decisions) for methods presented in this study, compared to evaluation of instances in random order. All our methods (PA, LSTM+AC, LSTM+DQN) mitigate bias and increase accuracy with respect to the ground-truth and outperform a heuristic mitigation procedure.

5.1 Retrospectively Mitigating Anchoring Bias with Probabilistic Adaptation (PA)

For college admission decisions, the PA strategy improves the accuracy to the ground-truth by around 2% in comparison to the decisions for a randomly ordered sequence. For the product review decisions, the accuracy to the ground-truth is improved by around 5% compared to the decisions for a randomly ordered sequence (Figure 5). The Pearson correlation coefficient, from here on referred to as a notion of "bias", decreases when using the PA strategy on both data sets. Using this strategy on college admissions, the bias decreases from -0.25 ($p < 0.01$) to -0.24 ($p < 0.01$). On the product review data set, the PA strategy decreases the correlation from -0.22 ($p < 0.01$) to -0.14 (p < 0.05). For negative ground-truth decisions, the accuracy stayed equal to the original data and improved by 7% for positive ground-truth for college admissions (Table 1a). For product reviews, the accuracy improved by 6% for negative ground-truth decisions and 9% for positive ground-truth decisions (Table 1b).

5.2 Prospectively Mitigating Anchoring Bias

Re-ordering the review sequences with the LSTM+RL procedures decreases bias and increases accuracy to the ground-truth. Both DQN and AC trained methods increase accuracy by over 7% compared to showing data instances randomly to evaluators. The bias decreases from -0.22 ($p < 0.01$) to -0.19 ($p < 0.01$) using DQN and to -0.15 ($p < 0.01$) using AC. A simple heuristic strategy improves the accuracy by 3%, and reduces the bias from -0.22 ($p < 0.01$) to -0.20 ($p < 0.01$). A comparison of all methods with respect to bias and accuracy is shown in Figure 5. The performance of our methods is analyzed in more detail for ground-truth positive and negative decisions in Table 1b. The true rejection accuracy is improved by 8% when showing files in a specific order learned by LSTM+DQN or LSTM+AC. This means that when presented with a negative review, an evaluator correctly classifies this review as negative in 8% more cases than when presented with the same instance in random order. When presented with a positive review, an evaluator also correctly marks the review as positive and indicates that they like the product in 18% more cases than when presented with the same instance in random order (Table 1b). The LSTM+DQN and LSTM+AC strategies were only evaluated on the low-stakes task of evaluating product

Algorithm	Ground-truth	Accuracy	Ground-truth	Accuracy
Random	-	0.84	+	0.90
PA	-	0.84	+	0.97
SVM	-	0.99	+	0.98

(a) College Admissions

Algorithm	Ground-truth	Accuracy	Ground-truth	Accuracy
Random	-	0.81	+	0.58
PA	-	0.87	+	0.67
Heuristic	-	0.84	+	0.74
SVM	-	0.76	+	0.99
LSTM+DQN	-	0.89	+	0.76
LSTM+AC	-	0.89	+	0.71

(b) Product Reviews

Table 1: Analysis of evaluator decisions made with our methods. We consider positive (+) and negative (-) ground-truth decisions separately. Our methods PA, LSTM+DQN and LSTM+AC improve the agreement to the ground-truth, which indicates that anchoring bias is mitigated.

reviews because experimentation in a high-stakes decision task like college admissions would raise ethical concerns.

6 DISCUSSION AND FUTURE WORK

Mitigating human biases is an important task in decision-making processes because biased human decisions can significantly impact people's lives. All proposed methods in our work mitigate incorrect decisions due to human anchoring bias and contribute to more fairness in decision-making.

6.1 The PA Mitigation Strategy has a larger impact on Product Preferences

For college admission decisions, there is a decrease of bias from -0.25 ($p < 0.01$) to -0.24 ($p < 0.01$), whereas for product reviews it decreases more significantly from -0.22 ($p < 0.01$) to -0.14 ($p < 0.05$). We believe this difference results from the training data that the SVM estimator is trained on. In college admissions, the SVM is trained with the final admission outcomes by the admissions committee. The final admission decision for a student depends on the admission recommendations of the individual evaluators. It is difficult to say how much this dependency influences the results, but we see a different outcome for product reviews. The SVM for the product reviews is trained on the rating of the review, which is independent of the decision of the evaluator.

We also observe that the PA strategy detects anchored decisions similarly well for both positive and negative ground-truth instances for product review decisions. For college admissions, the magnitude of detected negatively biased decisions is smaller. It would be interesting to evaluate whether negatively biased decisions are harder to detect in high-stakes decisions (e.g. college admissions) compared to low-stakes decisions like indicating interest in a product. We would like to see how to adapt the algorithm to bridge the gap between false positives and false negatives in future work.

6.2 The Importance of Model Prediction Performance

We observe that the SVM models used in our study outperform human prediction performance. We assume that the enhanced prediction capacities of the SVM's are important when aiming to adapt a human decision or re-sample the review sequence in case of a detected bias. This is because 1) the SVM prediction is used as a feature for the LSTM+RL procedure to learn the anchoring state of a reviewer and 2) the SVM prediction is directly used in the logical post-hoc adjustment in Equation 2. We can evaluate the performance of the model before including it in any decision-making process and subsequently decide if our models are suitable for the use case. In case we have a scenario where we can not train an ML model with sufficient accuracy, we can still use the probability of bias obtained from the exponential decay probability function from Equation 1 to find potentially biased decisions.

6.3 Complementary Team Performance for Mitigation of Human Anchoring Bias

It would be interesting to see if reviewers are less prone to being anchored when exposed to their anchoring state indicated by the exponential decay anchoring probability proposed in Section 3.1. This could be shown to evaluators as complementary information in an AI-Human team setup. Potentially, when made aware of their behavior, anchoring bias could be mitigated [4, 11]. Additionally, in case of an indication of anchoring, SVM predictions could be shown to an evaluator to show them a different perspective. Due to the strong influence of AI predictions on human decision-making [5, 37, 42], it might be possible to mitigate anchoring that way. However, this could also introduce other cognitive biases (e.g. confirmation or automation bias) [1, 3, 36] that might be hard to distinguish or dilute the attribution of bias and we leave this exploration to future work.

6.4 Enhancing Explainability of Anchoring Bias and Mitigation Strategies

Our results point to a significant impact of instance ordering on human cognitive bias, and we show that they can be mitigated with the help of ML. However, capturing anchoring bias with an LSTM model is based on black-box ML, and hence there is limited interpretability of the anchoring state produced by the LSTM network. Understanding the embedding space of the LSTM better could lead to more interpretability of the anchoring state [25]. A more interpretable model for capturing anchoring is our PA method, which quantifies the anchoring state of an evaluator with a probability. This probability is a metric on a human interpretable scale that can be shown to an evaluator to increase awareness of bias [11]. Future work could also analyze the relationship between the anchoring state and the next instance to be shown, because it is difficult to tell which strategy is used by DQN and AC to determine the order of instances shown to a reviewer. For this, frameworks like LIME [29] could be explored to form an understanding of algorithmic decisions and then used with concepts such as human-centered explainable AI [9, 31], visualization tools [17] or explainable AI with natural language [26]. The rules of which instance to show for which anchoring state could then be used by future researchers in their study design or be displayed to the user to increase awareness as a mitigation procedure.

6.5 Risk of Using Our Bias Mitigation Strategies

To evaluate the performance of our PA strategy, we adapt the decisions when an evaluator appears to be biased and measure the change of agreement to the ground-truth and the change in sequential dependencies with the correlation coefficient. Instead of adapting the human decisions, the probabilistic strategy could be used to indicate when an evaluator was potentially biased during the sequential decision process, and flag the respective instances to be reviewed again. Those flagged instances could either be reviewed again by different evaluators to establish consensus on the decision, be presented to the same evaluator, or be consolidated by a meta-reviewer. Future work could analyze if any unintentional biases are introduced when instances are flagged to be reviewed again.

We show that prospectively mitigating anchoring bias with parametric models helps evaluators make more unbiased decisions. Since this procedure does not show any additional information

or alters the information that decision-makers use to form their decision, this method does not impose additional risk compared to the established practice of showing instances in an arbitrary order, and the final decision still relies only on the human. With our procedure, the evaluator is not subject to a change in the task setup that could potentially introduce additional bias [4, 11]. We argue that our method is hence usable in a variety of sequential decision-making tasks.

7 CONCLUSION

Our study finds that it is possible to combine the strengths of AI systems and humans to enable fairer decisions in sequential decision-making tasks. Based on data of sequential college admission and rejection decisions at a large US university, and decisions on product preferences obtained in a controlled study on Amazon Mechanical Turk, we found evaluators to be biased by their previous decisions. Our study proposes mitigation strategies to detect and balance anchoring bias in those sequential decision tasks. We find that we can mitigate bias retrospectively for already-made decisions by capturing the anchoring state of a reviewer with a probabilistic model and adapting decisions with a logical strategy. This algorithm increases agreement to ground-truth by 2-5% and reduces sequential dependencies, or "bias", to previously made decisions (measured with the Pearson correlation coefficient) by 0.01-0.08. This model can additionally be used to flag instances to be re-reviewed, in case an evaluator was anchored by previously made decisions. We also show that we can learn in which order to present instances to an evaluator such that an evaluator is less biased by their own previous decisions. This prospectively mitigates bias before decisions are made. Simple heuristic re-sampling, like alternating strong and weak instances, does help to increase agreement by 3% and mitigate bias by 0.02, but learning the specific anchoring state of an evaluator leads to significantly fewer biased decisions. We reach an increased agreement to ground-truth by 7% and reduced bias of 0.07 using machine learning to learn an evaluator's anchoring state and learn the most favorable order in which to surface instances that minimizes anchored decisions. Our work has implications on individual fairness because it reinforces individuals with similar characteristics to be treated similarly, independent of when they were reviewed in a sequential decision-making process.

REFERENCES

[1] Gediminas Adomavicius, Jesse Bockstedt, Shawn Curley, and Jingjing Zhang. 2011. Recommender systems, consumer preferences, and anchoring effects. In *RecSys 2011 Workshop on Human Decision Making in Recommender Systems*. Citeseer, ACM, New York, NY, 35–42.

[2] Matias Barenstein. 2019. ProPublica's COMPAS Data Revisited.

[3] Nazlı Bhatia and Brian C Gunia. 2018. "I was going to offer $10,000 but...": The effects of phantom anchors in negotiation. *Organizational Behavior and Human Decision Processes* 148 (2018), 70–86.

[4] Richard A Block and David R Harper. 1991. Overconfidence in estimation: Testing the anchoring-and-adjustment hypothesis. *Organizational behavior and human decision processes* 49, 2 (1991), 188–207.

[5] Carrie J Cai, Samantha Winter, David Steiner, Lauren Wilcox, and Michael Terry. 2019. "Hello AI": Uncovering the Onboarding Needs of Medical Practitioners for Human-AI Collaborative Decision-Making. *Proceedings of the ACM on Human-computer Interaction* 3, CSCW (2019), 1–24.

[6] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. 2019. 23 ways to nudge: A review of technology-mediated nudging in human-computer interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–15.

[7] Andy Cockburn, Philip Quinn, and Carl Gutwin. 2015. Examining the peak-end effects of subjective experience. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, New York, NY, 357–366.

[8] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness through awareness. , 214–226 pages.

[9] Upol Ehsan and Mark O Riedl. 2020. Human-centered explainable ai: Towards a reflective sociotechnical approach. In *International Conference on Human-Computer Interaction*. Springer, Springer, New York, NY, 449–466.

[10] Christopher G. Harris. 2020. Mitigating Cognitive Biases in Machine Learning Algorithms for Decision Making. In *Companion Proceedings of the Web Conference 2020*. ACM, New York, NY, 775–781.

[11] Joey F George, Kevin Duffy, and Manju Ahuja. 2000. Countering the anchoring and adjustment bias with decision support systems. *Decision Support Systems* 29, 2 (2000), 195–206.

[12] Jianrui Huang, Xianyou He, Xiaojin Ma, Yian Ren, Tingting Zhao, Xin Zeng, Han Li, and Yiheng Chen. 2018. Sequential biases on subjective judgments: Evidence from face attractiveness and ringtone agreeableness judgment. *Plos one* 13, 6 (2018), e0198723.

[13] Peter J Huber. 1965. A robust version of the probability ratio test. *The Annals of Mathematical Statistics* 1, 1 (1965), 1753–1758.

[14] Walt Jesteadt, R Duncan Luce, and David M Green. 1977. Sequential effects in judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance* 3, 1 (1977), 92.

[15] Daniel Kahneman. 2011. *Thinking, Fast and Slow*. Macmillan, New York, NY.

[16] Yvonne Kammerer and Peter Gerjets. 2014. The role of search result position and source trustworthiness in the selection of web search results when using a list or a grid interface. *International Journal of Human-Computer Interaction* 30, 3 (2014), 177–191.

[17] Harmanpreet Kaur, Harsha Nori, Samuel Jenkins, Rich Caruana, Hanna Wallach, and Jennifer Wortman Vaughan. 2020. Interpreting Interpretability: Understanding Data Scientists' Use of Interpretability Tools for Machine Learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–14.

[18] Thomas C Leonard. 2008. Richard H. Thaler, Cass R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness.

[19] Weiwen Leung, Zheng Zhang, Daviti Jibuti, Jinhao Zhao, Maximilian Klein, Casey Pierce, Lionel Robert, and Haiyi Zhu. 2020. Race, Gender and Beauty: The Effect of Information Provision on Online Hiring Biases. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–11.

[20] Ariel Levy, Monica Agrawal, Arvind Satyanarayan, and David Sontag. 2021. Assessing the Impact of Automated Suggestions on Decision Making: Domain Experts Mediate Model Errors but Take Less Initiative. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–13.

[21] Gabriel Lima, Nina Grgić-Hlača, and Meeyoung Cha. 2021. Human Perceptions on Moral Responsibility of AI: A Case Study in AI-Assisted Bail Decision-Making. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–17.

[22] Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. 2016. Asynchronous methods for deep reinforcement learning. , 1928–1937 pages.

[23] Feng Ni, David Arnott, and Shijia Gao. 2019. The anchoring effect in business intelligence supported decision-making. *Journal of Decision Systems* 28, 2 (2019), 67–81.

[24] Jianmo Ni, Jiacheng Li, and Julian McAuley. 2019. Justifying recommendations using distantly-labeled reviews and fine-grained aspects. , 188–197 pages.

[25] Matteo Palmonari and Pasquale Minervini. 2020. Knowledge graph embeddings and explainable AI. *Knowledge Graphs for Explainable Artificial Intelligence: Foundations, Applications and Challenges, IOS Press, Amsterdam* 47, 1 (2020), 49–72.

[26] Nazneen Fatema Rajani, Bryan McCann, Caiming Xiong, and Richard Socher. 2019. Explain yourself! leveraging language models for commonsense reasoning.

[27] Juan Ramos et al. 2003. Using tf-idf to determine word relevance in document queries. , 29–48 pages.

[28] Charvi Rastogi, Yunfeng Zhang, Dennis Wei, Kush R Varshney, Amit Dhurandhar, and Richard Tomsett. 2020. Deciding Fast and Slow: The Role of Cognitive Biases in AI-assisted Decision-making.

[29] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*. ACM, New York, NY, 1135–1144.

[30] Melrose Roderick, James MacGlashan, and Stefanie Tellex. 2017. Implementing the deep q-network.

[31] Andrew Ross, Nina Chen, Elisa Zhao Hang, Elena L. Glassman, and Finale Doshi-Velez. 2021. Evaluating the Interpretability of Generative Models by Interactive Reconstruction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 80, 15 pages. https://doi.org/10.1145/3411764.3445296

[32] Sashank Santhanam, Alireza Karduni, and Samira Shaikh. 2020. Studying the effects of cognitive biases in evaluation of conversational agents. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–13.
[33] Linda J Skitka, Kathleen L Mosier, and Mark Burdick. 1999. Does automation bias decision-making? *International Journal of Human-Computer Studies* 51, 5 (1999), 991–1006.
[34] Tatsuji Takahashi, Masahiro Nakano, and Shuji Shinohara. 2010. Cognitive symmetry: Illogical but rational biases. *Symmetry: culture and science* 21, 1-3 (2010), 275–294.
[35] Poorna Talkad Sukumar, Ronald Metoyer, and Shuai He. 2018. Making a pecan pie: Understanding and supporting the holistic review process in admissions. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–22.
[36] Amos Tversky and Daniel Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *science* 185, 4157 (1974), 1124–1131.
[37] Karel van den Bosch and Adelbert Bronkhorst. 2018. Human-AI cooperation to benefit military decision making.
[38] Patrick Van Esch, J Stewart Black, and Joseph Ferolie. 2019. Marketing AI recruitment: The next phase in job application and selection. *Computers in Human Behavior* 90 (2019), 215–222.
[39] David W Vinson, Rick Dale, and Michael N Jones. 2019. Decision contamination in the wild: Sequential dependencies in online review ratings. *Behavior research methods* 51, 4 (2019), 1477–1484.
[40] Bing Xu, Naiyan Wang, Tianqi Chen, and Mu Li. 2015. Empirical evaluation of rectified activations in convolutional network.
[41] Qian Yang, John Zimmerman, Aaron Steinfeld, Lisa Carey, and James F Antaki. 2016. Investigating the heart pump implant decision process: opportunities for decision support tools to help. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 4477–4488.
[42] Rui Zhang, Nathan J McNeese, Guo Freeman, and Geoff Musick. 2021. " An Ideal Human" Expectations of AI Teammates in Human-AI Teaming. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–25.
[43] Yunfeng Zhang, Rachel KE Bellamy, and Wendy A Kellogg. 2015. Designing information for remediating cognitive biases in decision-making. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, New York, NY, 2211–2220.

8 APPENDIX

8.1 Technical Details on How to Use RL for Bias Mitgation

Resampling the Review Sequence with Actor-Critic (AC) Reinforcement Learning. To learn the order of instances to be shown to an evaluator that minimizes anchoring bias by previous decisions, we use an AC approach [22]. This algorithm consists of two neural networks – the actor and the critic. The actor decides on the next instance to be sampled from un-reviewed instances. The next possible instances are all instances of the review sequence which have not previously been surfaced to an evaluator. The critic then decides on the utility of the surfaced instance. In our LSTM+RL procedures, the probability of impossible actions, such as already surfaced instances, is set to a small constant probability to calculate the loss and update the network weights. It is set to zero at inference time, so that every instance can be shown only once. In our setup, the actor and the critic are both 3-layer multi-layer perceptrons. The hidden layers use a ReLU activation function [40]. The last layer for the actor produces probabilities for each action using a softmax activation. The critic produces a scalar output from its last linear layer to evaluate the utility of the instance. The two networks are updated with advantage actor-critic losses [22]. The action space was transformed from a continuous to a discrete representation due to the nature of being able to select a discrete number of instances to be shown next.

Resampling the Review Sequence with Deep-Q Learning. We use a second strategy to determine the optimal order of a review sequence using a Deep-Q Network (DQN) approach [30]. This algorithm uses a policy and a target network consisting of 3-layer multi-layer perceptrons with ReLU activation for the hidden layers and softmax for the final layer. The policy network decides on the next instance to surface, similar to the actor in the previous section. The anchoring state, surfaced instance, subsequent anchoring state, and reward associated with this transition are saved to replay memory. The network is updated by sampling transitions from replay memory and calculating the smooth L1 loss [13] between the state-action (sampled instance) values and expected state-action values. This is the discounted expected value of the next sampled instance plus the current reward.

The goal is to minimize the anchoring state of an evaluator. This goal is measured in the reward for updating the RL models $r_t = min \sum_{j=1}^{n} 1 - h_j$, where h is the anchoring state obtained from the LSTM (Section 3.3). We can consider $a_{t_0}, a_{t_1}, \ldots, a_{t_i} \in \mathbb{A}$, where $\mathbb{A}$ is a set of possible "actions", or data instances to be shown to the evaluator for a review sequence with length i (e.g. sequentially evaluating i students in a row). Algorithm 1 shows the pseudocode for our LSTM+RL learning procedures. It learns a policy mapping $\pi : \mathbb{S} \rightarrow \mathbb{A}$ from anchoring states $h_{t_0}, h_{t_1}, \ldots, h_{t_i} \in \mathbb{S}$ (where $\mathbb{S}$ are all possible anchoring states) to instance files, which should be shown to a human evaluator. These procedures *learn how to mitigate anchoring states by sampling specific kinds of instances to re-balance the anchoring state of the evaluator.*

Algorithm 1: Pseudocode to mitigate anchoring.

Result: $\pi : \mathbb{S} \rightarrow \mathbb{A}$
foreach *review sequence* **do**
  for *data instance index i* **do**
    $\hat{d}_{t_0,\ldots,t_i}, h_{t_0,\ldots,t_i} = LSTM(d_{t_0,\ldots,t_{i-1}})$ *# simulated decisions and anchor state for* $t_0 - t_i$;
    $p_{s_t} = \pi(h_{t_0,\ldots,t_i})$*# probabilities to decide on next instance,* $a_t \in \mathbb{A}$;
    $r_t = \sum_{j=0}^{n} 1 - h_j$*# reward*;
    $\mathcal{L}_{DQN}$ =*[13]# DQN loss* ;
    $\mathcal{L}_{AC}$ =*[22]# AC loss*;
    # backpropagate and update weights for policy π;
  end
end