

# Online Conformal Selection with Accept-to-Reject Changes

Kangdao Liu<sup>1\*</sup>, Huajun Xi<sup>2\*</sup>, Chi-Man Vong<sup>1†</sup>, Hongxin Wei<sup>2†</sup>

<sup>1</sup>Department of Computer and Information Science, University of Macau

<sup>2</sup>Department of Statistics and Data Science, Southern University of Science and Technology  
kangdaoliu@gmail.com, 12112806@mail.sustech.edu.cn, cmvong@um.edu.mo, weihx@sustech.edu.cn

## Abstract

Selecting a subset of promising candidates from a large pool is crucial across various scientific and real-world applications. *Conformal selection* offers a distribution-free and model-agnostic framework for candidate selection with uncertainty quantification. While effective in offline settings, its application to online scenarios, where data arrives sequentially, poses challenges. Notably, conformal selection permits the deselection of previously selected candidates, which is incompatible with applications requiring irreversible selection decisions. This limitation is particularly evident in resource-intensive sequential processes, such as drug discovery, where advancing a compound to subsequent stages renders reversal impractical. To address this issue, we extend conformal selection to an online *Accept-to-Reject Changes* (ARC) procedure: non-selected data points can be reconsidered for selection later, and once a candidate is selected, the decision is irreversible. Specifically, we propose a novel conformal selection method, *Online Conformal Selection with Accept-to-Reject Changes* (dubbed **OCS-ARC**), which incorporates online Benjamini–Hochberg procedure into the candidate selection process. We provide theoretical guarantees that OCS-ARC controls the false discovery rate (FDR) at or below the nominal level at any timestep under both i.i.d. and exchangeable data assumptions. Additionally, we theoretically show that our approach naturally extends to multivariate response settings. Extensive experiments on synthetic and real-world datasets demonstrate that OCS-ARC significantly improves selection power over the baseline while maintaining valid FDR control across all examined timesteps.

## Introduction

Selecting a subset of promising candidates from a large pool is crucial in various scientific and real-world applications, such as recruitment screening (Heaslip 2022), drug discovery (Sheridan et al. 2015), and foundation model alignment (Gui, Jin, and Ren 2024). For instance, in recruitment, companies identify qualified applicants to enhance organizational performance and outcomes. This gives rise to the significance of *conformal selection* (Jin and Candès 2023a,b), a systematic framework for candidate selection with uncertainty quantification. Specifically, conformal selection constructs a con-

formal  $p$ -value to quantify whether a candidate should be selected and applies the Benjamini-Hochberg (BH) procedure (Benjamini and Hochberg 1995; Benjamini and Yekutieli 2001) to control the False Discovery Rate (FDR) with a theoretical guarantee. Importantly, this framework imposes no parametric assumptions about the data distribution and is applicable to any black-box predictor, making it a trustworthy and powerful technique for candidate selection.

While conformal selection excels in offline settings, its application to online scenarios, where data arrives sequentially, poses challenges. In particular, when applied to this setting, conformal selection allows previously selected candidates to be deselected later. However, such reversibility may be incompatible with some real-world applications:

- In drug discovery, due to high costs and time constraints, compound selection often occurs in sequential stages. Once a compound is selected at any timestep, it may advance to subsequent experiments, rendering reversal at later points impractical. However, non-selected candidates can be reconsidered later, ensuring that significant compounds are not permanently disregarded.
- In recruitment screening, applications are received on a rolling basis and evaluated sequentially in batches until all required positions are filled. Once a candidate is identified as qualified and selected, they may advance to subsequent evaluations (e.g., interviews), with withdrawals at later stages being uncommon. Non-selected candidates may be placed on a waitlist for potential reconsideration.

In such cases, a selection mechanism that satisfies an Accept-to-Reject Changes (ARC) procedure is required: non-selected data points can be reconsidered for selection in a later stage, but once a candidate is selected, the decision is rendered final and irreversible. Consequently, extending conformal selection to this setting is a critical step toward enabling trustworthy candidate selection in online scenarios.

In this work, we extend conformal selection to the online ARC setting. We propose a novel method, *Online Conformal Selection with Accept-to-Reject Changes* (**OCS-ARC**), which satisfies the ARC property. In particular, OCS-ARC incorporates the online Benjamini-Hochberg procedure (Fischer, Xu, and Ramdas 2024) into the selection of conformal  $p$ -values. We provide a theoretical guarantee for OCS-ARC, demonstrating that, under the assumption of independent and

\*These authors contributed equally.

†Correspond to Chi-Man Vong and Hongxin Wei.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

identically distributed (i.i.d.) data, the method controls the FDR at or below the nominal level at any timestep. This guarantee also extends to more general settings where the data satisfies a natural exchangeability condition. Moreover, we demonstrate that OCS-ARC can be naturally extended to multivariate response settings. In particular, we formally establish that our method maintains valid FDR control in such scenarios, as long as the non-conformity functions employed by OCS-ARC satisfy the *regional monotonicity* property.

To verify the validity and effectiveness of OCS-ARC, we conduct extensive experiments to evaluate its performance in terms of FDR and power across both synthetic datasets and real-world applications. For real-world scenarios, we apply OCS-ARC to recruitment screening, drug property prediction, and LLM decision-making within the conformal alignment framework (Gui, Jin, and Ren 2024). As this is the first work to extend conformal selection to the online ARC setting, no existing baseline is available for direct comparison. To address this, we introduce a baseline by adapting the Bonferroni correction to the online ARC setting. The results show that our proposed method significantly improves the selection power over the baseline, while ensuring valid FDR at each timestep, across all scenarios. In practice, our approach is easy to implement with any predictive algorithm, such as deep models, and incurs negligible computational costs.

## Preliminary

**Problem Setup.** We study the problem of candidate selection in *online* setting where the test data arrives in a sequential order (Foster and Stine 2008; Javanmard and Montanari 2018; Xu and Ramdas 2024; Fischer, Boffill Roig, and Brannath 2024; Fischer, Xu, and Ramdas 2024). Formally, define the feature space  $\mathcal{X} \subseteq \mathbb{R}^d$  and response space  $\mathcal{Y} \subseteq \mathbb{R}$ . We sample a sequence of *test* data points  $\{(\mathbf{X}_{n+t}, Y_{n+t})\}$ ,  $t \in \mathbb{N}^+$ , from the joint distribution  $\mathcal{P}_{\mathcal{X}\mathcal{Y}}$ , where the outcomes  $\{Y_{n+t}\}$  are unobserved. At each timestep  $t$ , we observe a new instance  $\mathbf{X}_{n+t}$  and receive a pre-defined threshold  $c_t$ . Our goal is to *identify a subset of indices*  $\mathcal{R}_t \subseteq [t]$  *from the test data set that maximizes the number of test observations*  $i \in \mathcal{R}_t$  *satisfying*  $Y_i > c_i$ . For notation shorthand, we denote  $[t] = \{1, \dots, t\}$ .

To provide a theoretical guarantee, the selection procedure is required to control the FDR (Benjamini and Hochberg 1995) at each timestep  $t$  below a user-specified level  $q$ . The FDR is defined as the expected proportion of false discoveries (i.e.  $i \in \mathcal{R}_t$  but  $Y_i \leq c_i$ ) among all selected observations:

$$\text{FDR}_t = \mathbb{E} \left[ \frac{|\mathcal{R}_t \cap \mathcal{H}_t^0|}{|\mathcal{R}_t|} \right],$$

where  $\mathcal{H}_t^0 = \{i \in [t] : Y_i \leq c_i\}$ , with  $0/0 = 0$ .

To evaluate selection performance, we calculate the expected proportion of correct selections, referred to as *power*:

$$\text{Power}_t = \mathbb{E} \left[ \frac{|\mathcal{R}_t \cap \mathcal{H}_t^1|}{|\mathcal{H}_t^1|} \right],$$

where  $\mathcal{H}_t^1 = \{i \in [t] : Y_i > c_i\}$ . An ideal selection procedure should maximize  $\text{Power}_t$  while controlling the  $\text{FDR}_t$  at or below the user-specified nominal level for each timestep  $t$ .

**Conformal Selection.** *Conformal selection* (Jin and Candès 2023b) is a distribution-free and model-agnostic selection framework that provides theoretical guarantees for finite-sample FDR control. To gather statistical evidence for selection, conformal selection partitions a subset of the training set as the calibration set. In the following, we denote calibration set as  $\mathcal{D}_{\text{cal}} = \{(\mathbf{X}_i, Y_i)\}_{i=1}^n$  and test set as  $\mathcal{D}_{\text{test}} = \{(\mathbf{X}_{n+t}, Y_{n+t})\}_{t=1}^T$ . Conformal selection addresses the following multiple hypothesis testing problem:

$$H_t^0 : Y_{n+t} \leq c_t \quad \text{vs.} \quad H_t^1 : Y_{n+t} > c_t, \quad \forall t = 1, \dots, m.$$

Rejecting the null hypothesis  $H_t^0$  indicates the  $t$ -th test sample should be included in the selection set, i.e.,  $t \in \mathcal{R}$ , as its response is deemed to exceed the given threshold  $c_t$ .

To construct the selection set  $\mathcal{R}$ , conformal selection employs *conformal p-value*, which builds upon the conformal inference framework (Vovk, Gammerman, and Saunders 1999; Vovk, Gammerman, and Shafer 2005). Given the pre-trained model, we define a *non-conformity score* function  $V : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ . The non-conformity score measures how well a data point  $(\mathbf{X}, Y)$  aligns with the calibration data, with a larger non-conformity score suggesting that the point is more likely to be an outlier relative to the calibration set. For calibration samples, the non-conformity scores are  $V_i = V(\mathbf{X}_i, Y_i)$ , and for test samples, they are  $\hat{V}_{n+t} = V(\mathbf{X}_{n+t}, c_t)$ , where  $c_t$  is the pre-defined threshold that replaces the unobserved  $Y_{n+t}$ . These scores are then used to compute conformal  $p$ -values through a rank-based comparison of  $\hat{V}_{n+t}$  against the calibration scores  $\{V_i\}_{i=1}^n$ . A smaller conformal  $p$ -value indicates a lower rank of  $\hat{V}_{n+t}$  relative to the calibration scores. This provides stronger evidence for rejecting  $H_t^0$ , i.e., selecting the  $t$ -th test sample.

To determine the final selection set  $\mathcal{R}$ , conformal selection employs the BH procedure (Benjamini and Hochberg 1995; Benjamini and Yekutieli 2001), a widely adopted approach for controlling the FDR in multiple hypothesis testing, to select conformal  $p$ -values. The BH procedure guarantees that the overall FDR is maintained at or below the user-specified level, expressed as  $\text{FDR} \leq q$ . A comprehensive explanation of the BH procedure is provided in Appendix A.

**Accept-to-Reject Changes Procedure.** In real-world applications, such as recruitment screening and drug discovery, a selection mechanism that adheres to the *Accept-to-Reject Changes* (ARC) procedure is crucial.

**Definition 1.** An Accept-to-Reject Changes procedure is a nested sequence of selection sets:

$$\mathcal{R}_1 \subseteq \mathcal{R}_2 \subseteq \dots \subseteq \mathcal{R}_t \subseteq \dots,$$

where  $\mathcal{R}_t$  denotes the selection set at timestep  $t$ .

**Remark 2.** While the term ‘Accept-to-Reject Changes’ might superficially suggest a process where selected items can later be rejected, this interpretation is misleading. The term originates from the language of hypothesis testing, where ‘rejecting’ a null hypothesis corresponds to selecting a data point. Thus, ‘Accept-to-Reject Changes’ refers to the process that permits transitions from acceptance (non-selection) to rejection (selection), while prohibiting the reverse.

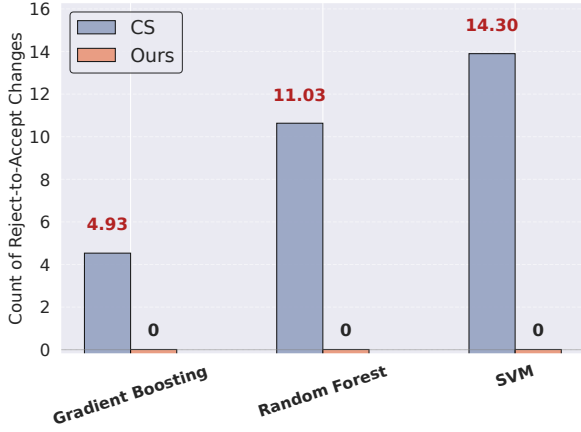


Figure 1: Comparison of conformal selection (CS) with OCS-ARC. This experiment measures the total number of reject-to-accept samples for both methods on 500 sequentially arriving data points from a synthetic dataset across three models. Results are averaged over 30 independent runs.

An ARC procedure allows non-selected data points to be reconsidered for selection later, and once a point is selected, the decision is final and irreversible. However, conformal selection is not an ARC procedure: when applying it at different timesteps, it would change its earlier selection into deselection. This can be illustrated by the following example:

**Example 3.** Consider a FDR level of  $q = 0.1$ , with conformal  $p$ -values given as  $p_1 = 0.1$  and  $p_2 = 0.2$ . Applying BH at timestep 1 results in  $\mathcal{R}_1 = \{1\}$ . At the second timestep, reapplying the procedure yields the selection set  $\mathcal{R}_2 = \emptyset$ .

This example illustrates that conformal selection may alter the status of a candidate from being selected to being deselected. To empirically validate it, we conduct synthetic experiments across various models and record the number of data points that transition from being selected to deselected (i.e., reject-to-accept). Detailed setups are provided in Appendix B. The results in Figure 1 confirm that this framework violates the ARC property. *This motivates our method, which extends conformal selection to an online ARC setting.*

## Method

In this section, we introduce our method, *Online Conformal Selection with Accept-to-Reject Changes* (OCS-ARC). Our approach is composed of three primary steps:

- Training:** Train a predictive model  $\hat{\mu}$  for  $Y$ . The model can be constructed using any appropriate algorithm.
- Calibration:** Construct a monotone non-conformity function using the model  $\hat{\mu}$  and evaluate it on both the calibration set and the test set. Using the resulting scores, calculate the conformal  $p$ -values for each test sample.
- Thresholding:** At each timestep  $t$ , apply the online BH procedure (Fischer, Xu, and Ramdas 2024) to the set of conformal  $p$ -values, yielding the selection set  $\mathcal{R}_t$ .

Both the training and calibration steps rely on the labeled set  $\mathcal{D}_{\text{train}}$ . In the case where a predictive model  $\hat{\mu}$  is already

available, our OCS-ARC can be directly applied using the entire labeled set  $\mathcal{D}_{\text{train}}$  for calibration. Otherwise, the training data is split into two subsets: one for model training and the other for calibration. For simplicity, in the following discussion, we assume that the model  $\hat{\mu}$  is already available (pre-trained), and all labeled data are utilized for calibration. We denote the calibration set as  $\mathcal{D}_{\text{cal}} = \{(\mathbf{X}_i, Y_i)\}_{i=1}^n$  and the test set as  $\mathcal{D}_{\text{test}} = \{(\mathbf{X}_{n+t}, Y_{n+t})\}_{t=1}^T$ . In our setting, *the calibration set is fully available beforehand, whereas the test data arrive sequentially*. This setup is common in prior works (Bao et al. 2024a; Sale and Ramdas 2025).

We next provide a detailed explanation of the *calibration* process. Specifically, we employ conformal  $p$ -values (Bates et al. 2023; Jin and Candès 2023b) to perform the hypothesis testing. When the true responses  $\{Y_{n+t}\}_{t=1}^T$  are observed, we compute the non-conformity score  $V_{n+t} = V(\mathbf{X}_{n+t}, Y_{n+t})$ . Subsequently, the  $t$ -th *oracle* conformal  $p$ -value is defined as

$$\frac{\sum_{i=1}^n \mathbb{1}\{V_i < V_{n+t}\} + U_t(1 + \sum_{i=1}^n \mathbb{1}\{V_i = V_{n+t}\})}{n + 1},$$

where  $U_t \sim \text{Uniform}[0, 1]$  are i.i.d uniform random variable to randomize over ties when  $V_{n+t}$  equals some  $V_i$ , ensuring a continuous conformal  $p$ -value. Standard results from conformal inference ensure that if the data sample  $(\mathbf{X}_{n+t}, Y_{n+t})$  follows the same distribution as the calibration data, then the oracle conformal  $p$ -value  $p_t^*$  stochastically dominates the uniform distribution on  $[0, 1]$  (Bates et al. 2023):

$$\mathbb{P}\{p_t^* \leq \alpha\} \leq \alpha, \quad \forall \alpha \in [0, 1]. \quad (1)$$

This property is essential to ensure FDR control. However, the computation of the oracle conformal  $p$ -value is *infeasible*, as the true responses  $\{Y_t\}_{j=1}^m$  are unobserved. Therefore, we replace  $V_{n+t}$  with its estimate  $\hat{V}_{n+t} = V(\mathbf{X}_{n+t}, c_t)$ , resulting in the  $t$ -th (practical) conformal  $p$ -values:

$$\frac{\sum_{i=1}^n \mathbb{1}\{V_i < \hat{V}_{n+t}\} + U_t(1 + \sum_{i=1}^n \mathbb{1}\{V_i = \hat{V}_{n+t}\})}{n + 1}. \quad (2)$$

Here,  $p_j$  quantifies how extreme the threshold is compared to the usual behavior of the outcomes, with a smaller  $p_j$  providing stronger evidence for rejecting  $H_j^0$ . To ensure that the practical conformal  $p$ -value  $p_j$  stochastically dominates the uniform distribution on  $[0, 1]$  (see Eq. (1)), the non-conformity function  $V$  is required to satisfy the following *monotonicity* property (Jin and Candès 2023a,b):

**Definition 4.** A non-conformity score function  $V(\cdot, \cdot) : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$  is monotone if  $V(\mathbf{X}, y) \leq V(\mathbf{X}, y')$  holds for any  $\mathbf{X} \in \mathcal{X}$  and any  $y, y' \in \mathcal{Y}$  obeying  $y \leq y'$ .

In the Experiments section, we will showcase several examples of monotone non-conformity score functions. After calculating the practical conformal  $p$ -values, we apply the online BH procedure (Fischer, Xu, and Ramdas 2024), which leverages a sequence of real numbers  $\{\gamma_t\}_{t=1}^T$  satisfying  $\sum_{t=1}^T \gamma_t \leq 1$  for all  $T \in \mathbb{N}^+$ . The complete OCS-ARC procedure is formally outlined in Algorithm 1.

**Theoretical guarantee.** We now provide theoretical guarantees for OCS-ARC. First, in the following theorem, we establish that OCS-ARC satisfies the ARC procedure.

---

**Algorithm 1: Online Conformal Selection with Accept-to-Reject Changes (OCS-ARC)**


---

**Require:** Calibration data  $\mathcal{D}_{\text{cal}} = \{(\mathbf{X}_i, Y_i)\}_{i=1}^n$ , test data  $\mathcal{D}_{\text{test}} = \{(\mathbf{X}_{n+t}, Y_{n+t})\}_{t=1}^T$ , predefined threshold  $\{c_t\}_{t=1}^T$ , FDR nominal level  $q \in (0, 1)$ , non-conformity score  $V : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ , real numbers  $\{\gamma_t\}_{t=1}^T$ .

- 1: Compute  $V_i = V(\mathbf{X}_i, Y_i)$  for  $i = 1, \dots, n$ .
- 2: **for**  $t = 1, \dots, T$  **do**
- 3:   Compute  $\hat{V}_{n+t} = V(\mathbf{X}_{n+t}, c_t)$ .
- 4:   Construct conformal  $p$ -value  $p_t$  as in Eq. (2)
- 5:   Compute online BH procedure threshold:

$$k_t^* = \max\{k \in [t] : \sum_{j=1}^t \mathbb{1}\{p_j \leq k\alpha\gamma_j\} \geq k\}$$

- 6:   Construct selection set:  $\mathcal{R}_t = \{j \in [t] : p_j \leq k_t^* q \gamma_j\}$
- 7: **end for**

---

**Theorem 5.** *OCS-ARC is an ARC procedure. In particular, the selection sets  $\{\mathcal{R}_t\}_{t=1}^T$  outputted by Algorithm 1 satisfy*

$$\mathcal{R}_1 \subseteq \mathcal{R}_2 \subseteq \dots \subseteq \mathcal{R}_T.$$

The proof is provided in Appendix C.1. To validate our theoretical findings, we conduct synthetic experiments on various regression models and count the number of points transitioning from selected to deselected (reject-to-accept). Detailed descriptions of the experimental setups are provided in Appendix B. The results, shown in Figure 1, confirm that OCS-ARC satisfies the properties of an online ARC procedure. Then, we show that Algorithm 1 controls the FDR at any timestep  $t$ , under the assumption that the calibration and test data are identically and independently distributed.

**Theorem 6.** *Consider a monotone non-conformity score function  $V : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ . Assume that the calibration data  $\{(\mathbf{X}_i, Y_i)\}_{i=1}^n$  and the test data  $\{(\mathbf{X}_{n+t}, Y_{n+t})\}_{t=1}^T$  are independently and identically distributed. Then, for any nominal level  $q \in (0, 1)$  and timestep  $t$ , the selection set  $\mathcal{R}_t$  constructed by Algorithm 1 satisfies  $\text{FDR}_t \leq q$ .*

The detailed proof of Theorem 6 is provided in Appendix C.2. In the following theorem, we extend our previous result to the scenario where the calibration and test samples satisfy a natural exchangeability condition. The corresponding technical proof is provided in Appendix C.3.

**Theorem 7.** *Consider a monotone non-conformity score function  $V$ . Assume that  $\{V_1, \dots, V_n, V_{n+t}\}$  are exchangeable conditional on  $\{\hat{V}_{n+t'}\}_{t'=1, t' \neq t}^T$  and  $\{V_{t'}\}_{t'=1}^{n+T}$  have no ties almost surely. Then, for any nominal level  $q \in (0, 1)$  and timestep  $t$ , the corresponding selection set  $\mathcal{R}_t$  constructed by our OCS-ARC in Algorithm 1 satisfies  $\text{FDR}_t \leq q$ .*

**Remark 8.** It is important to emphasize that our method guarantees FDR control only under the i.i.d. or exchangeability assumption. This guarantee does not hold in dynamic environments with distribution shifts, as seen in online learning. The key distinction lies in the availability of feedback. In online learning, after making a decision, the model receives

feedback and adapts its policy accordingly. In contrast, our setting lacks such feedback: once a selection is made, no information about its correctness is provided.

**Remark 9.** We highlight two key theoretical novelties in OCS-ARC. In standard online BH procedure, the  $p$ -value are assumed to satisfy conditional validity:

$$\mathbb{P}\{p_t \leq \alpha \mid t \in \mathcal{H}_t^0\} \leq \alpha.$$

However, conformal  $p$ -value only ensures marginal validity:

$$\mathbb{P}\{p_t \leq \alpha, t \in \mathcal{H}_t^0\} \leq \alpha.$$

Moreover, online BH procedure relies on the assumption that  $p$ -values are *positive regression dependence on a subset* (PRDS) to guarantee FDR control. Nevertheless, conformal  $p$ -value fails to meet this requirement. Due to these issues, directly applying online BH to conformal  $p$ -value is nontrivial.

**Extension to multivariate conformal selection.** Recent work (Bai et al. 2025) extends conformal selection to multivariate response settings. In this section, we show that our OCS-ARC procedure can also be applied in this context. Specifically, for a  $d'$ -dimensional multivariate response  $\mathbf{Y}_{n+t}$ , consider the following hypothesis testing:

$$H_t^0 : \mathbf{Y}_{n+t} \in R^c \quad \text{vs.} \quad H_t^1 : \mathbf{Y}_{n+t} \in R, \quad \forall t = 1, \dots, m,$$

where  $R$  represents an arbitrary pre-defined closed target region in  $\mathbb{R}^{d'}$ . To construct a selection set, we define a new non-conformity score function  $V(\cdot, \cdot) : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ . We compute  $V_i = V(\mathbf{X}_i, \mathbf{Y}_i)$ , for  $i = 1, \dots, n$ , and  $\hat{V}_{n+t} = V(\mathbf{X}_{n+t}, \mathbf{r}_t)$ , for  $t = 1, \dots, T$ , where  $\mathbf{r}_t$  is an arbitrary closed point in target region  $R$ . Then, we compute the conformal  $p$ -values defined in Eq. (2). To guarantee that the conformal  $p$ -value  $p_t$  stochastically dominates the uniform distribution on  $[0, 1]$ , the non-conformity function  $V$  must satisfy *regional monotonicity* property (Bai et al. 2025):

**Definition 10.** A non-conformity score function  $V(\cdot, \cdot) : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$  is regional monotone if  $V(X, y) \leq V(X, y')$  holds for any  $X \in \mathcal{X}$  and any  $y \in R^c, y' \in R$ .

Examples of regional monotone non-conformity score functions are provided in Appendix D. After computing the conformal  $p$ -values, the online BH procedure (Fischer, Xu, and Ramdas 2024) is applied to obtain the final selection set. The whole procedure for *Multivariate Online Conformal Selection with Accept-to-Reject Changes* (dubbed mOCS-ARC) is summarized in Appendix D. In the following theorem, we show that mOCS-ARC controls FDR at any timestep  $t$ .

**Theorem 11.** *Consider a regional monotone non-conformity score function  $V : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ . Assume that the calibration data  $\{(\mathbf{X}_i, \mathbf{Y}_i)\}_{i=1}^n$  and the test data  $\{(\mathbf{X}_{n+t}, \mathbf{Y}_{n+t})\}_{t=1}^T$  are independently and identically distributed. Then, for any nominal level  $q \in (0, 1)$  and timestep  $t$ , the selection set  $\mathcal{R}_t$  constructed by mOCS-ARC satisfies  $\text{FDR}_t \leq q$ .*

The technical proof of this theorem is provided in Appendix C.4. This theoretical result establishes that our approach naturally extends to the multivariate response setting.

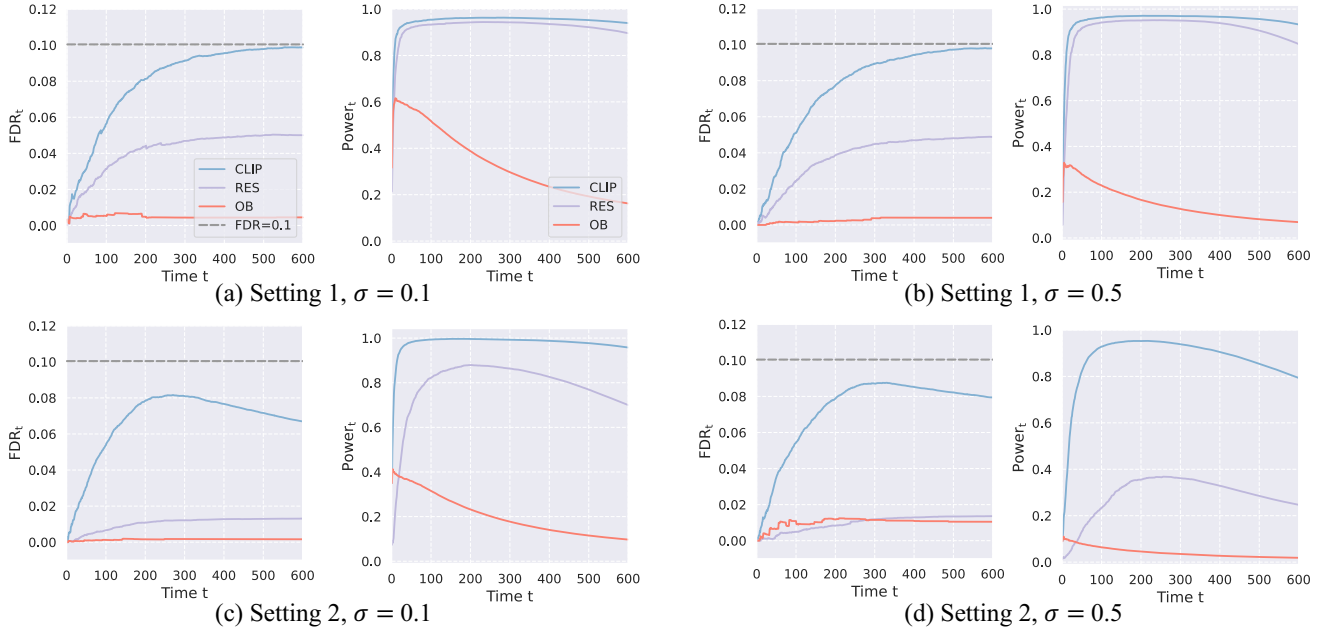


Figure 2: Synthetic data experiments with varying noise levels across different data-generating processes. The target FDR is set to 0.10, and gradient boosting is used to fit the regression model across all configurations. Subplots (a) and (b) present results for simulation setting 1 (CLIP, RES, and OB, respectively), while subplots (c) and (d) correspond to simulation setting 2.

## Experiments

We conduct a comprehensive evaluation of OCS-ARC in terms of FDR and power across synthetic datasets and real-world applications. For simulated experiments, we follow the setup in prior work (Jin and Candès 2023b). In real-world scenarios, we apply OCS-ARC to candidate screening, drug property prediction, and LLM decision-making within the conformal alignment framework (Gui, Jin, and Ren 2024). *Notably, the experimental setups and results for drug property prediction, along with the relevant mOCS-ARC experiments, are detailed in Appendices F.2 and D, respectively.* The source code and supplementary appendix are publicly available at <https://github.com/J4ckLiu/OCS-ARC>.

### Numerical Experiments

**Setup.** We evaluate our method in simulated datasets for regression. The task is to select individuals with responses  $Y_{n+t} > 0$  from sequentially arrived test data while controlling FDR at each timestep  $t$ . We generate i.i.d. covariates  $\mathbf{X}_i \sim \text{Uniform}[-1, 1]^{20}$  and responses  $\mathbf{Y}_i = \mu(\mathbf{X}_i) + \epsilon_i$ , where  $\mu(\mathbf{x}) = \mathbb{E}[Y | \mathbf{X} = \mathbf{x}]$  is nonlinear in  $\mathbf{x}$ , and  $\epsilon_i$  represents random noise independent of  $\mathbf{X}$ . To demonstrate the performance of our methods under different data-generating processes, we apply two simulation settings:

- **Setting 1:**  $\mu(\mathbf{x}) = 4x_1 \cdot \mathbf{1}(x_2 > 0) \cdot \max(0.5, x_3) + 4x_1 \cdot \mathbf{1}(x_2 \leq 0) \cdot \min(-0.5, x_3)$ ,
- **Setting 2:**  $\mu(\mathbf{x}) = 5x_1x_2 + e^{x_4-1}$ .

In both settings, the noise is modeled as  $\epsilon_i \sim N(0, \sigma^2)$  with homogeneous variance. We adjust  $\sigma$  to different levels to generate data with varying noise levels. The sizes of the training

and calibration datasets are fixed at  $|\mathcal{D}_{\text{train}}| = |\mathcal{D}_{\text{calib}}| = 1000$ , while the size of the test dataset is set to  $|\mathcal{D}_{\text{test}}| = 600$ .

We utilize gradient boosting and support vector machines (SVM) with a radial basis function kernel to fit a regression model, using the scikit-learn Python library. We apply Algorithm 1 with two different monotone non-conformity score functions, following prior work (Jin and Candès 2023b):

- CLIP:  $V(x, y) = 1000 \cdot \mathbf{1}\{y > 0\} - \hat{\mu}(x)$
- RES:  $V(x, y) = y - \hat{\mu}(x)$ .

For brevity, we refer to OCS-ARC with CLIP as CLIP, and OCS-ARC with RES as RES. For  $\{\gamma_t\}_{t=1}^T$ , we set

$$\gamma_t = r^t \cdot \frac{1-r}{r},$$

where  $r$  represents a predefined decay coefficient. In our main experiments,  $r$  is fixed at 0.99 without additional tuning.

**Baseline.** To establish a comparison, we introduce a baseline by adapting the Bonferroni correction to the online ARC setting, referred to as *Online Bonferroni correction with Accept-to-Reject Changes* (dubbed OB).

OB: Select all  $p_t \leq q\gamma_t$  with score function CLIP.

OB employs the same  $\{\gamma_t\}_{t=1}^T$  as in OCS-ARC. We show that OB satisfies an ARC procedure in Appendix E.

**Results.** The results for gradient boosting are presented in Figure 2, while the results for SVM are included in the Appendix F.1. For the evaluation of  $\text{FDR}_t$ , we average the  $\text{FDP}_t$  over 300 independent runs at different timesteps. The results show that the proposed method achieves an  $\text{FDR}_t$  close to the nominal level 0.1, whereas the baseline exhibits a larger

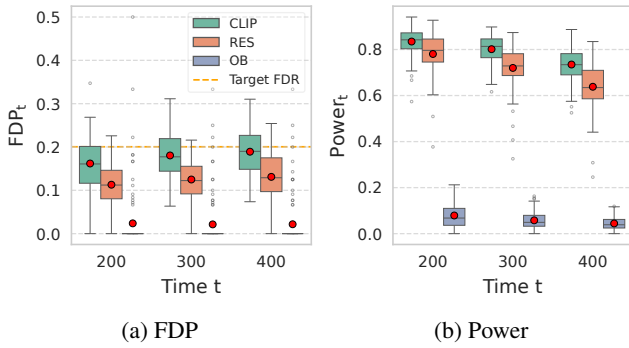


Figure 3: Results for candidate screening in recruitment. The target FDR is set to 0.20, and gradient boosting is used to fit the classifier.  $FDP_t$  and  $Power_t$  are reported over 100 independent runs at timesteps 100, 200 and 300.

deviation from it. Among the two non-conformity score functions, CLIP consistently demonstrates a higher realized  $FDP_t$  across all settings, remaining closer to 0.1. Additionally, the  $FDR_t$  stays below 0.1 in all configurations. For  $Power_t$ , we evaluate it by calculating the average proportion of correct selections among all positive samples up to timestep  $t$  across all replicates. We observe that our methods consistently outperform the baseline OB under all scores and simulation settings. Moreover, CLIP consistently achieves higher  $Power_t$ ; however, the power decreases as noise strength increases across all settings. This decline arises because increased noise impairs the model accuracy, thereby deteriorating the selection performance. The results demonstrate that OCS-ARC surpasses the baseline, significantly improving selection power.

### Experiments on Recruitment Screening

We apply OCS-ARC as an automated screening tool for recruitment, where HR staff use machine learning predictions to sequentially evaluate incoming candidates. Applicants may be either shortlisted for subsequent tests and interviews or placed on a waiting list. Here, the model estimates the likelihood that a candidate is qualified, with higher scores indicating a better fit. We employ OCS-ARC to calibrate predictions and produce a shortlist with rigorous FDR control.

**Setup.** We utilize a recruitment dataset available on Kaggle (El Kharoua 2024), which contains 1,500 entries. We divide the dataset into three parts: 700 samples for training, 400 for calibration, and the remaining 400 for testing. A gradient boosting classifier is employed to perform predictions for this task. We aim to select as many data points with labels of 1 as possible while controlling the FDR rate at a predefined level. Furthermore, the target FDR is set to 0.2 for this experiment, and CLIP is applied as the non-conformity score function.

**Results.** In Figure 3, we use  $FDP_t$  at  $t = 100, 200, 300$  over 100 replications to evaluate  $FDR_t$ . The results show that the proposed method achieves an  $FDR_t$  close to the nominal level 0.2, while OB exhibits a large deviation. For  $Power_t$ , we average the proportion of correct selections among all positive samples up to the corresponding timesteps over all

replicates. We observe that our methods consistently outperform OB across all scores. Overall, the proposed method significantly improves the selection power over the baseline, while ensuring valid FDR control at each timestep.

### Application to Large Language Models

Given the widespread application of LLMs in decision-making, we show that OCS-ARC can be integrated with LLMs within the conformal alignment framework (Gui, Jin, and Ren 2024). In this experiment, we aim to assess a stream of question-answers and select high-quality answers while controlling the FDR. Further details on the conformal alignment framework are provided in Appendix G. Notably, this procedure can be readily extended to *any* other online candidate selection task utilizing LLMs as base models.

**Setup.** The experiment follows the default settings of conformal alignment (Gui, Jin, and Ren 2024), employing a logistic regression model as the alignment score predictor. We utilize two widely recognized question-answering datasets: TriviaQA (Joshi et al. 2017) and CoQA (Reddy, Chen, and Manning 2019). We employ self-evaluation likelihood, input uncertainty scores, and output confidence scores as features to train the alignment score predictor. The experiments are conducted using LLaMA-2-13B-chat (Touvron et al. 2023), with the target FDR varied from 0 to 0.80 in fixed increments over 25 steps. For each dataset, we randomly select 1,000 samples, including 400 for training the alignment score predictor, 400 for calibration, and 200 samples as test data.

**Results.** The experimental results on TriviaQA are presented in Figure 4, while the results on CoQA are provided in Appendix F.3. We empirically evaluate the  $FDP_t$  at different timesteps  $t = 50, 100, 150, 200$  over 100 replications. Unlike prior experiments, we examine the performance of OCS-ARC at different target FDRs. Results show that our method controls the average  $FDP_t$  below each target FDR at every evaluated timestep. Moreover, our method achieves high  $Power_t$ , indicating that it effectively selects most of the high-quality answers. Overall, OCS-ARC offers a model-agnostic candidate selection framework that can be effortlessly integrated with any black-box model, including LLMs.

### Parameter Sensitivity Analysis

In this section, we investigate the effect of  $r$  and the size of the calibration set (SoC) on the performance of OCS-ARC.

**Results.** We set the target FDR at 0.1, employ gradient boosting for model fitting, and generate data based on Setting 1 from our main numerical experiments. Figure 5a illustrates the impact of varying the decay coefficient  $r$ , selected from the set  $\{0.99, 0.993, 0.996, 0.999\}$ . The results show that our method consistently ensures valid FDR control across different  $r$ . Moreover, a larger  $r$  (e.g., 0.999) typically leads to a more conservative approach at smaller  $t$ , resulting in a relatively lower  $Power_t$ . In Figure 5b, we demonstrate the performance of our method under different calibration set sizes  $\{1000, 2000, 3000\}$ , at  $t = 100, 200, 300$ . We observe that a larger calibration set slightly reduces the variance of FDR and power, leading to more stable outcomes. In general,

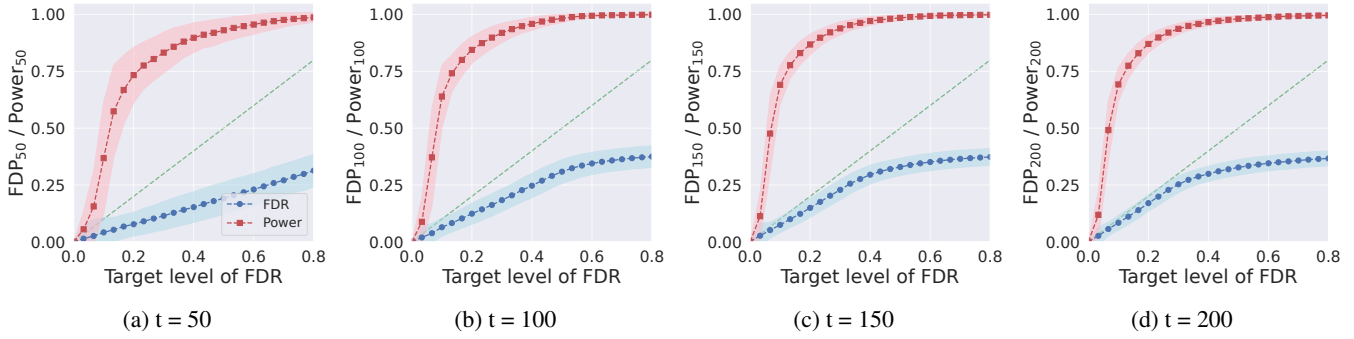


Figure 4: Results for the application of OCS-ARC to question answering with LLMs, on the TriviaQA Dataset. Experiments are conducted using LLaMA-2-13B-chat, with the target FDR varied from 0 to 0.80 in fixed steps over 25 increments.  $FDP_t$  and  $Power_t$  are reported over 100 independent runs at timesteps 50, 100, 150, and 200.

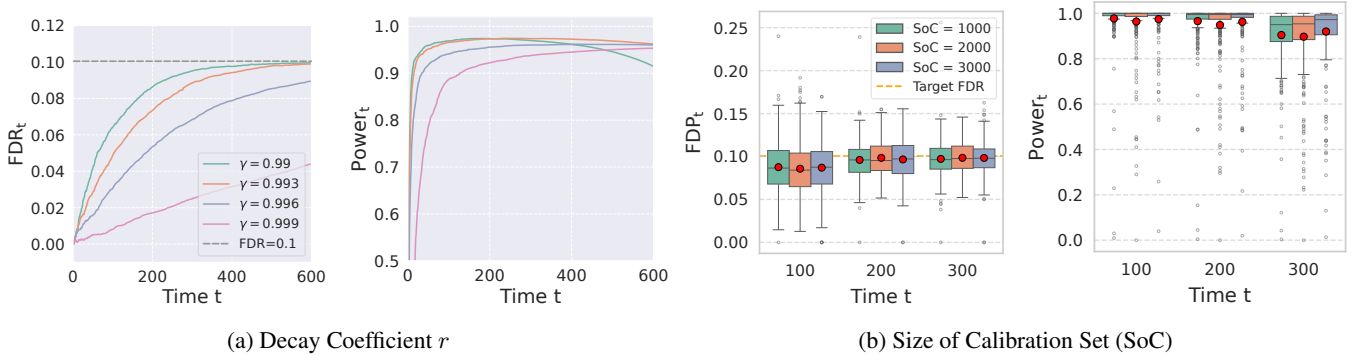


Figure 5: Results of parameter sensitivity analyses, including (a) the effect of varying the decay coefficient  $r$ , and (b) the effect of changing the size of the calibration set. The target FDR is fixed at 0.10, and gradient boosting is used to fit the regression model. Parameter  $r$  is chosen from the set  $\{0.99, 0.993, 0.996, 0.999\}$ , while SoC is varied over  $\{1000, 2000, 3000\}$ .

our method is robust to changes in the calibration set size. Overall, the results confirm the robustness of our method with respect to its parameters and the calibration set size.

### Related Work

Conformal inference (Papadopoulos et al. 2002; Vovk, Gamerman, and Shafer 2005) is an emerging technique for uncertainty quantification that constructs prediction intervals guaranteed to contain the true label with a predefined probability. It has been applied across domains, including regression (Lei and Wasserman 2014; Romano, Sesia, and Candès 2020; Zeng et al. 2025), classification (Angelopoulos et al. 2021; Huang et al. 2024; Xi et al. 2025a; Zhou et al. 2025; Liu et al. 2025), and online settings (Gibbs and Candès 2021; Bhatnagar et al. 2023; Angelopoulos, Barber, and Bates 2024; Xi et al. 2025b). Within the framework, several studies focus on controlling the FDR in predictive settings, known as *conformalized multiple testing* (Bates et al. 2023; Jin and Candès 2023b; Hu and Lei 2024; Zhang et al. 2025). This encompasses two-sample tests (Hu and Lei 2024) and multiple testing applications such as outlier detection (Zhang et al. 2022; Bates et al. 2023; Zhang et al. 2025) and data selection/sampling (Jin and Candès 2023a,b). A recent line of work extends this framework to selective conformal inference

that constructs prediction intervals for a specific subset of test data, under both offline (Bao et al. 2024b; Jin and Ren 2025) and online settings (Bao et al. 2024a; Sale and Ramdas 2025). However, these approaches do not address the *online candidate selection* problem, where selection decisions are irrevocable—a critical requirement in practical applications like recruitment screening. In this work, we extend conformal selection to the online ARC setting to bridge this gap.

### Conclusion

In this work, we propose OCS-ARC, an extension of conformal selection to the online ARC setting. We provide theoretical guarantees for OCS-ARC, showing that the method controls the FDR at or below the nominal level at any timestep under both i.i.d. and exchangeable data assumptions. Furthermore, we theoretically demonstrate that our approach can be easily extended to multivariate response settings. Extensive experiments show that our method significantly improves the selection power over the baseline, while ensuring valid FDR control at each timestep. We hope the insights of this work can enhance the applicability and impact of conformal selection across diverse scientific and industrial domains.

## Acknowledgments

This research is supported by the Shenzhen Science and Technology Innovation Committee (Project No. SGDX20220530111001006), the Macau Science and Development Fund (Grants No. 0118/2024/RIA2 and 0216/2024/AGJ), and the Shenzhen Fundamental Research Program (Grant No. JCYJ20230807091809020).

We gratefully acknowledge the support of the Center for Computational Science and Engineering at the Southern University of Science and Technology for our research.

## References

- Angelopoulos, A. N.; Barber, R.; and Bates, S. 2024. Online conformal prediction with decaying step sizes. In *Forty-first International Conference on Machine Learning*.
- Angelopoulos, A. N.; Bates, S.; Jordan, M. I.; and Malik, J. 2021. Uncertainty Sets for Image Classifiers using Conformal Prediction. In *9th International Conference on Learning Representations*.
- Bai, T.; Zhao, Y.; Yu, X.; and Yang, A. Y. 2025. Multivariate conformal selection. *International Conference of Machine Learning*.
- Bao, Y.; Huo, Y.; Ren, H.; and Zou, C. 2024a. Cap: A general algorithm for online selective conformal prediction with fcr control. *arXiv preprint arXiv:2403.07728*.
- Bao, Y.; Huo, Y.; Ren, H.; and Zou, C. 2024b. Selective conformal inference with false coverage-statement rate control. *Biometrika*.
- Bates, S.; Candès, E.; Lei, L.; Romano, Y.; and Sesia, M. 2023. Testing for outliers with conformal p-values. *The Annals of Statistics*.
- Benjamini, Y.; and Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*.
- Benjamini, Y.; and Yekutieli, D. 2001. The control of the false discovery rate in multiple testing under dependency. *Annals of statistics*.
- Bhatnagar, A.; Wang, H.; Xiong, C.; and Bai, Y. 2023. Improved Online Conformal Prediction via Strongly Adaptive Online Learning. In *International Conference on Machine Learning*.
- El Kharoua, R. 2024. Predicting Hiring Decisions in Recruitment Data. Retrieved July 25, 2025, from <https://www.kaggle.com/datasets/rabieelkharoua/predicting-hiring-decisions-in-recruitment-data>.
- Fischer, L.; Bofill Roig, M.; and Brannath, W. 2024. The online closure principle. *The Annals of Statistics*.
- Fischer, L.; Xu, Z.; and Ramdas, A. 2024. Online generalizations of the e-BH and BH procedure. *arXiv preprint arXiv:2407.20683*.
- Foster, D. P.; and Stine, R. A. 2008.  $\alpha$ -investing: a procedure for sequential control of expected false discoveries. *Journal of the Royal Statistical Society Series B: Statistical Methodology*.
- Gibbs, I.; and Candès, E. 2021. Adaptive conformal inference under distribution shift. *Advances in Neural Information Processing Systems*.
- Gui, Y.; Jin, Y.; and Ren, Z. 2024. Conformal Alignment: Knowing When to Trust Foundation Models with Guarantees. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.
- Heaslip, E. 2022. Ai tools for talent acquisition to help you hire. <https://vervoe.com/ai-tools-for-talent-acquisition/>.
- Hu, X.; and Lei, J. 2024. A two-sample conditional distribution test using conformal prediction and weighted rank sum. *Journal of the American Statistical Association*, 119(546): 1136–1154.
- Huang, J.; Xi, H.; Zhang, L.; Yao, H.; Qiu, Y.; and Wei, H. 2024. Conformal Prediction for Deep Classifier via Label Ranking. In *Forty-first International Conference on Machine Learning*.
- Javanmard, A.; and Montanari, A. 2018. Online rules for control of false discovery rate and false discovery exceedance. *The Annals of statistics*.
- Jin, Y.; and Candès, E. J. 2023a. Model-free selective inference under covariate shift via weighted conformal p-values. *arXiv preprint arXiv:2307.09291*.
- Jin, Y.; and Candès, E. J. 2023b. Selection by prediction with conformal p-values. *Journal of Machine Learning Research*.
- Jin, Y.; and Ren, Z. 2025. Confidence on the focal: Conformal prediction with selection-conditional coverage. *Journal of the Royal Statistical Society Series B: Statistical Methodology*.
- Joshi, M.; Choi, E.; Weld, D. S.; and Zettlemoyer, L. 2017. Triviaqa: A large scale distantly supervised challenge dataset for reading comprehension. *arXiv preprint arXiv:1705.03551*.
- Lei, J.; and Wasserman, L. 2014. Distribution-free prediction bands for non-parametric regression. *Journal of the Royal Statistical Society Series B: Statistical Methodology*.
- Liu, K.; Sun, T.; Zeng, H.; Zhang, Y.; Pun, C.-M.; and Vong, C.-M. 2025. Spatial-aware conformal prediction for trustworthy hyperspectral image classification. *IEEE Transactions on Circuits and Systems for Video Technology*.
- Papadopoulos, H.; Proedrou, K.; Vovk, V.; and Gammernan, A. 2002. Inductive Confidence Machines for Regression. In Elomaa, T.; Mannila, H.; and Toivonen, H., eds., *Machine Learning: ECML 2002, 13th European Conference on Machine Learning*.
- Reddy, S.; Chen, D.; and Manning, C. D. 2019. Coqa: A conversational question answering challenge. *Transactions of the Association for Computational Linguistics*, 7: 249–266.
- Romano, Y.; Sesia, M.; and Candès, E. J. 2020. Classification with Valid and Adaptive Coverage. In Larochelle, H.; Ranzato, M.; Hadsell, R.; Balcan, M.; and Lin, H., eds., *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems*.
- Sale, Y.; and Ramdas, A. 2025. Online Selective Conformal Prediction: Errors and Solutions. *arXiv preprint arXiv:2503.16809*.

Sheridan, R. P.; McMasters, D. R.; Voigt, J. H.; and Wildey, M. J. 2015. eCounterscreening: using QSAR predictions to prioritize testing for off-target activities and setting the balance between benefit and risk. *Journal of Chemical Information and Modeling*.

Touvron, H.; Lavril, T.; Izacard, G.; Martinet, X.; Lachaux, M.-A.; Lacroix, T.; Rozière, B.; Goyal, N.; Hambro, E.; Azhar, F.; et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Vovk, V.; Gammerman, A.; and Saunders, C. 1999. Machine-learning applications of algorithmic randomness.

Vovk, V.; Gammerman, A.; and Shafer, G. 2005. *Algorithmic learning in a random world*. Springer.

Xi, H.; Huang, J.; Liu, K.; Feng, L.; and Wei, H. 2025a. Does confidence calibration improve conformal prediction? *Transactions on Machine Learning Research*.

Xi, H.; Liu, K.; Zeng, H.; Sun, W.; and Wei, H. 2025b. Robust online conformal prediction under uniform label noise. *arXiv preprint arXiv:2501.18363*.

Xu, Z.; and Ramdas, A. 2024. Online multiple testing with e-values. In *International Conference on Artificial Intelligence and Statistics*. PMLR.

Zeng, H.; Liu, K.; Jing, B.; and Wei, H. 2025. Parametric scaling law of tuning bias in conformal prediction. *International Conference of Machine Learning*.

Zhang, S.; Zhou, C.; Liu, Y.; Zhang, P.; Lin, X.; and Pan, S. 2025. Conformal Anomaly Detection in Event Sequences. In *Forty-second International Conference on Machine Learning*.

Zhang, Y.; Jiang, H.; Ren, H.; Zou, C.; and Dou, D. 2022. AutoMS: automatic model selection for novelty detection with error rate control. *Advances in Neural Information Processing Systems*, 35: 19917–19929.

Zhou, X.; Zeng, H.; Xia, X.; Jing, B.; and Wei, H. 2025. Semi-Supervised Conformal Prediction With Unlabeled Non-conformity Score. *arXiv preprint arXiv:2505.21147*.