CoCo: Conformal Confidence Suppression to Optimize Search Results

Zhou Qin[*]
Amazon
New York, NY, USA
qinzhouhit@hotmail.com

Shuang Wu[*]
Amazon
New York, NY, USA
shuangwu222@ucla.edu

Yoon Kim
Amazon
Palo Alto, CA, USA
yoonkim@amazon.com

Yi Liu
Amazon
Seattle, WA, USA
liuyi.feier@gmail.com

Wenyang Liu
Amazon
Seattle, WA, USA
wenyang6@ucla.edu

Abstract

Modern e-commerce recommendation systems aim to improve customer experience by ranking content on search results page (SRP). However, displaying content is not always beneficial for customers across all contexts; even top-ranked content can be irrelevant, misleading, or redundant in certain scenarios. In this work, we propose a robust content suppression mechanism to selectively suppress content when necessary. Our approach leverages causal effect learning to measure the incremental value of showing versus suppressing content. Prior work uses Conditional Average Treatment Effect (CATE) to estimate treatment effect without considering the inherent uncertainty in uplift predictions, resulting in over-suppression and degraded customer experience. Our work introduces a novel **Co**nformal **Co**nfidence (CoCo) suppressor that explicitly accounts for prediction uncertainty in uplift-based modeling. We first evaluate it on synthetic datasets with controlled noise, bias, and contexts. Results demonstrate superior performance compared to baseline approaches. Subsequently, our online traffic tests show statistically significant improvements in revenue and profit compared to existing methods.

CCS Concepts

• **Information systems** → **Recommender systems**; **Learning to rank**.

Keywords

Search Ranking; Conformal Prediction; Uplift Modeling

ACM Reference Format:
Zhou Qin, Shuang Wu, Yoon Kim, Yi Liu, and Wenyang Liu. 2026. CoCo: Conformal Confidence Suppression to Optimize Search Results. In *Proceedings of the 49th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '26), July 20–24, 2026, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3805712.3809847

1 Introduction

Modern recommendation and ranking systems aim to maximize user satisfaction and engagement by recommending or ranking content based on predicted relevance or reward [6]. Despite advances in uplift modeling for recommendation [3, 4], challenges persist in ensuring that the recommended content consistently improves user experience and meets business objectives. In particular, the presence of content with detrimental impacts—such as irrelevant, misleading, or low-quality recommendations—can pose a critical problem, necessitating robust content suppression mechanisms. For example, in Figure 1, we should suppress content A (e.g., A could be a themed widget containing video or bundled items) and move up the remaining search results if displaying A has a negative estimated value for the given search request.

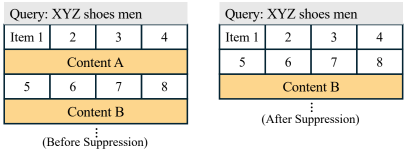

Figure 1: Suppression Illustration

One emerging approach to addressing this challenge is causal effect learning, which provides a framework to quantify the incremental value (uplift) of displaying versus suppressing a content. Without a suppressor, the ranking system simply displays top-K items regardless of quality. Prior work has leveraged Conditional Average Treatment Effect (CATE) for tasks such as personalized targeting, treatment effect estimation, and counterfactual reasoning in offline recommendation [2, 7]. However, these approaches often overlook the inherent uncertainty from uplift predictions, leading to over-suppression or under-suppression of content. Over-suppression excludes potentially beneficial content, while under-suppression fails to remove detrimental ones, both degrading system performance.

In the most common scenario, the decision to suppress content is based on whether the uplift value of displaying it compared to a baseline is negative—that is, content is suppressed when its presentation yields negative value to the user. A natural approach to constructing a baseline is to use average values, which, however, can fail to accurately represent the true counterfactual scenario. In practice, we observe that setting baselines using averages leads to excessively high suppression rates, necessitating manual intervention such as imposing a suppression rate cap (e.g., capping suppression at 5% of the times an item appears as a candidate).

---

[*] Both authors contributed equally to this research.

SIGIR '26, Melbourne, VIC, Australia

ACM ISBN 979-8-4007-2599-9/2026/07
https://doi.org/10.1145/3805712.3809847

These limitations underscore the need for more robust methods capable of accurately estimating the uplift value of content presentation with greater confidence. To bridge the gap, we adopt the Upper Confidence Bound (UCB) logic, suppressing content only when we are confident the uplift is negative, and introduce a novel Conformal Confidence (CoCo) suppressor that explicitly accounts for prediction uncertainty. Rather than training separate deep learning models to estimate uncertainty, which adds latency overhead and often requires parametric assumptions and lacks reliability under distribution shift, we leverage the conformal prediction framework [1] to calibrate these bounds. This allows for a model-agnostic, distribution-free estimate of uncertainty that is both theoretically rigorous and more computationally efficient than maintaining auxiliary uncertainty models. By incorporating model uncertainty, CoCo suppressors enable more informed suppression decisions, mitigating over-suppression while maintaining high precision and recall.

Our work makes the following contributions: (i) We propose a family of Conformal Confidence (CoCo) suppressors—CoCo-1, CoCo-M, and CoCo-F—that leverage conformal confidence bounds at varying levels of contextual granularity. CoCo-1 employs a single global threshold for suppression. CoCo-M utilizes multiple thresholds for different groups (e.g., grouping by region on the Search Results Page). CoCo-F applies a threshold function to enable personalized suppression decisions. (ii) These CoCo suppressors adaptively suppress content based on predicted uplift and associated uncertainty, enabling robust filtering even under noisy or biased predictions. (iii) We evaluate the proposed CoCo suppressors on synthetic datasets with controlled noise, bias, and contextual configurations. Results demonstrate that our methods achieve superior performance compared to baseline approaches, highlighting their robustness across diverse scenarios. (iv) We deploy this method in a production e-commerce environment and validate its effectiveness through an online A/B test, demonstrating improved performance over the existing production approach.

2 Methodology

2.1 Problem Formulation

We formulate this problem as causal effect learning within a search system's recommendation and ranking pipeline. Given a query instance $\boldsymbol{x}$ — which includes personalized user information and contextual information about the environment (e.g., membership, region) in addition to describing the query itself — let W denote the set of available rankable items. For each query feature instance $\boldsymbol{x}$, a ranker-recommending policy π ranks the items and recommends the top items to the user. For simplicity, the user's feedback is recorded as a reward y.

From this search recommendation process, we collect N samples as quadruples $\{(\boldsymbol{x}_i, w_i, a_i, y_i)\}_{i=1}^{N}$, where each tuple $(\boldsymbol{x}_i, w_i, a_i, y_i)$ represents: given query feature $\boldsymbol{x}_i$, a reward y_i is received when item $w_i \in \mathcal{W}$ is displayed (if $a_i = 1$) or withheld (if $a_i = 0$). In this setup, $(\boldsymbol{x}, w, a = 1, y)$ represents a factual sample, while $(\boldsymbol{x}, w, a = 0, y)$ represents a counterfactual sample for policy π. For simplicity, we define $\mathcal{D}$ to represent $\{(\boldsymbol{x}_i, w_i, y_i)\}_{i=1}^{N}$, with further distinctions for the factual dataset $\mathcal{D}_{a=1} := \{(\boldsymbol{x}_i, w_i, y_i)\}_{i:a_i=1}$ and counterfactual dataset $\mathcal{D}_{a=0} := \{(\boldsymbol{x}_i, w_i, y_i)\}_{i:a_i=0}$. We also define $\mathcal{D}_x := \{(\boldsymbol{x}_i, w_i)\}_{i=1}^{N}$ as the query-item data and use $\mathcal{D}(w)$, $\{\mathcal{D}_{a=a'}(w)\}_{a'=0,1}$, $\mathcal{D}_x(w)$ to represent the data with item w.

The expected uplift value for an item w given query $\boldsymbol{x}$ is defined as: $\Delta(\boldsymbol{x}, w) = \mathbb{E}[y|\boldsymbol{x}, w, a = 1] - \mathbb{E}[y|\boldsymbol{x}, w, a = 0]$. This expected uplift, also known as CATE, quantifies the incremental reward given query feature $\boldsymbol{x}$. A positive $\Delta(\boldsymbol{x}, w)$ indicates an increase in reward, while a negative value suggests a detrimental effect. CATE defines the target estimand, but in practice the predicted uplift $\hat{\Delta}(\boldsymbol{x}, w)$ carries prediction uncertainty. Suppression decisions based on point estimates alone lead to over- or under-suppression, which motivates our CoCo uncertainty-aware suppression framework.

2.2 Suppression

We design the suppression module to filter out items with negative expected uplift ($\Delta(\boldsymbol{x}, w) < 0$). Let $\hat{\Delta}(\boldsymbol{x}, w)$ denote the predicted uplift for item w given a query $\boldsymbol{x}$. A straightforward approach would be to suppress item w if $\hat{\Delta}(\boldsymbol{x}, w) < 0$. However, this direct approach neglects the uncertainty in model prediction, leading to "over-suppression" or "under-suppression". Specifically, suppose an uplift $\Delta(\boldsymbol{x}, w)$ is positive while our prediction $\hat{\Delta}(\boldsymbol{x}, w)$ is negative due to noise in the training data or randomness from model training, then this suppression logic will suppress item w given the query feature $\boldsymbol{x}$, which blocks the potential reward gain from item w.

Specifically, to address the "over-suppression" problem, we aim to allow more impressions if we lack sufficient confidence that the uplift is truly negative. In practice, over-suppression is the more prevalent issue, because point-estimate uplift predictions that neglect uncertainty tend to suppress items whose true uplift is positive but whose predicted uplift is negative due to noise, thereby removing potentially beneficial content and causing revenue drop. To achieve this, we set up a hypothesis test with a null hypothesis $H_0 : \Delta(\boldsymbol{x}, w) \geq 0$ and an alternative hypothesis $H_a : \Delta(\boldsymbol{x}, w) < 0$. Our suppression logic relies on a one-sided confidence interval $(-\infty, \hat{\Delta}(\boldsymbol{x}, w) - \mathrm{T}(\boldsymbol{x}, w)]$: we suppress item w only if its associated upper confidence bound is negative (i.e., $\hat{\Delta}(\boldsymbol{x}, w) < \mathrm{T}(\boldsymbol{x}, w)$, where T denotes the threshold). This approach introduces a tolerance threshold, relaxing suppression for items with uncertain predictions. For under-suppression, a similar test can be set up with $H_0 : \Delta(\boldsymbol{x}, w) \leq 0$ and $H_a : \Delta(\boldsymbol{x}, w) > 0$ using a lower confidence bound. In the remainder of this paper, we focus on handling over-suppression since under-suppression involves a similar hypothesis test and confidence bound construction.

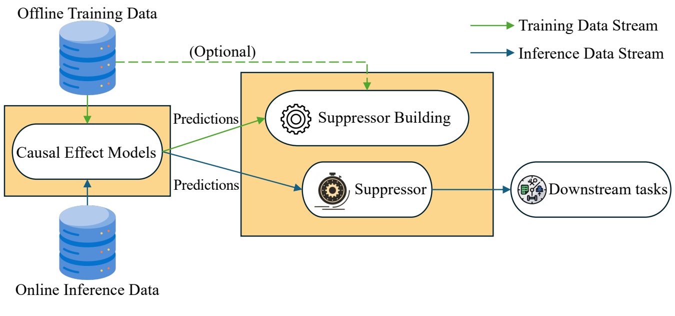

Figure 2: Suppression Module: causal effect models trained offline, applied online for downstream tasks.

T-Learners. Our methodology begins with T-Learners, a meta-learner approach that trains two separate models for causal effect estimation [5]. We employ two separate deep-learning-based models, one for factual data $f_{a=1}$, estimating $\mathbb{E}[y|\boldsymbol{x}, w, a = 1]$, and another for counterfactual data $f_{a=0}$, estimating $\mathbb{E}[y|\boldsymbol{x}, w, a = 0]$ (i.e., models in Figure 2). The predicted uplift is then computed as:

$$\hat{\Delta}(\boldsymbol{x}, w) = f_{a=1}(\boldsymbol{x}, w) - f_{a=0}(\boldsymbol{x}, w). \tag{1}$$

These two models, $f_{a=1}$ and $f_{a=0}$, can be based on any model and are trained on the factual dataset $\mathcal{D}_{a=1}$ and counterfactual dataset $\mathcal{D}_{a=0}$, respectively. To train each model, we use the Mean Squared Error (MSE) loss for continuous rewards and Cross-Entropy (CE) loss for binary rewards. Once we have trained models, we can then construct a suppressor to determine the suppression threshold and make suppression decisions during online inference, i.e., deciding not to show items based on estimated uplift value and relativity with suppression threshold.

2.3 Conformal Confidence Suppressor

We propose constructing conformal confidence bound in suppression module, motivated by conformal prediction technique [1]. We propose our method as CoCo suppressor module. This technique suggests a way to construct prediction intervals that are valid under minimal assumptions about the data distribution. It provides a rigorous interval estimate for uplift prediction $\Delta(\boldsymbol{x}, w)$, ensuring a specified confidence level.

As introduced in Section 2.2, our goal is to determine an upper confidence bound $\hat{\Delta}(\boldsymbol{x}, w) - \mathrm{T}(\boldsymbol{x}, w)$, where the threshold $\mathrm{T}(\boldsymbol{x}, w)$ is proportional to the standard deviation estimate $\hat{\sigma}(\boldsymbol{x}, w)$. This threshold incorporates uncertainty in the prediction $\hat{\Delta}(\boldsymbol{x}, w)$, quantified by standard deviation. We face two primary challenges: (i) estimating the model uncertainty $\hat{\sigma}(\boldsymbol{x}, w)$ and (ii) determining if $\hat{\sigma}(\boldsymbol{x}, w)$ should be item-specific or based on coarser granularity, such as a more general group of items. Accordingly, we develop three types of CoCo suppressors, CoCo-1, CoCo-M and CoCo-F and the suppressors are introduced below.

2.3.1 CoCo-1. When using a single threshold in our suppressor, we propose two versions of conformal-confidence-based suppressor, called CoCo-1-α and CoCo-1-β where 1 represents the single threshold output. In this part of CoCo-1, for simplicity, we denote the true uplift, predicted uplifts and observed uplift data as $\{\Delta_i\}_{i=1}^n$, $\{\hat{\Delta}_i\}_{i=1}^n$, $\{\delta_i\}_{i=1}^n$ respectively. In both versions, the CoCo suppressor performs suppression: $\hat{\Delta}(\boldsymbol{x}, w) < \mathrm{T}$ at coverage level $1 - \alpha$.

CoCo-1-α. In the α version, we use the actual residuals as the nonconformity scores and use the quantile from the scores as the threshold. The steps are as follows: (i) Compute residuals as the nonconformity scores: $s_i = \hat{\Delta}_i - \delta_i$. (ii) Find the sample α-th quantile for scores $\{s_i\}_{i=1}^n$. (iii) Output q_α as T.

CoCo-1-β. In the β version, we use deviations of model predictions from the sample mean as the nonconformity scores and use the quantile from the scores as the threshold. Steps are as follows (i) Compute sample deviations as nonconformity scores: $s_i = \hat{\Delta}_i - \bar{\Delta}$ where $\bar{\Delta}$ is the sample mean of $\{\hat{\Delta}_i\}_{i=1}^n$. (ii) Find the sample α-th quantile q_α for scores $\{s_i\}_{i=1}^n$. (iii) Output q_α as T.

2.3.2 CoCo-M. We also propose a conformal confidence based suppressor for multiple thresholds grouped by some variable, called CoCo-M, where M means multiple groups, e.g., different regions of the SRP can have different thresholds. Denote the contextual variable as $g(\boldsymbol{x}, w)$, which is a function of feature $\boldsymbol{x}$ and item w. Assuming $g(\boldsymbol{x}, w)$ takes K groups $g_1, \ldots, g_K$, CoCo-M then outputs K thresholds. In both α and β versions, the CoCo suppressor performs suppression: $\hat{\Delta}(\boldsymbol{x}, w) < \mathrm{T}(\boldsymbol{x}, w)$ with confidence $1 - \alpha$.

CoCo-M-α. In the α version, we use normalized actual residual, grouped by contextual variable, as the nonconformity scores and use the quantile from the scores as the threshold. Steps are as follows: (i) Compute residuals $r(\boldsymbol{x}_i, w_i) = \hat{\Delta}(\boldsymbol{x}_i, w_i) - \delta(\boldsymbol{x}_i, w_i)$. (ii) Compute sample means $\{\bar{r}(g_k)\}_{k\in[K]}$ and sample standard deviations $\{\hat{\sigma}(g_k)\}_{k\in[K]}$ for all groups. (iii) Compute normalized residuals as the nonconformity scores: $s(\boldsymbol{x}_i, w_i) = \frac{\hat{\Delta}(\boldsymbol{x}_i, w_i) - \delta(\boldsymbol{x}_i, w_i)}{\hat{\sigma}(g(\boldsymbol{x}_i, w_i))}$. (iv) Find the sample α-th quantile for scores $\{s(\boldsymbol{x}_i, w_i)\}_{i=1}^n$. (v) Output q_α and $\mathrm{T}(\boldsymbol{x}_i, w_i) := q_\alpha \cdot \hat{\sigma}(g(\boldsymbol{x}_i, w_i))$, $\forall i \in [n]$.

CoCo-M-β. In the β version, we use sample deviations which are normalized by group standard deviation, as the nonconformity scores and use the quantile from the scores as the threshold. Steps are as follows. (i) Compute sample mean $\{\bar{\Delta}(g_k)\}_{k\in[K]}$ and sample standard deviations $\{\hat{\sigma}(g_k)\}_{k\in[K]}$ for all groups. (ii) Compute normalized sample deviations as the nonconformity scores: $s(\boldsymbol{x}_i, w_i) = \frac{\hat{\Delta}(\boldsymbol{x}_i, w_i) - \bar{\Delta}(g(\boldsymbol{x}_i, w_i))}{\hat{\sigma}(g(\boldsymbol{x}_i, w_i))}$. (iii) Find the sample α-th quantile for scores $\{s(\boldsymbol{x}_i, w_i)\}_{i=1}^n$. (iv) Output $\{q_\alpha \cdot \hat{\sigma}(g_k)\}_{k\in[K]}$ as $\mathrm{T}(\boldsymbol{x}, w)$.

2.3.3 CoCo-F. CoCo-F is a robust approach to assess model uncertainty involving interval prediction, rather than point-wise prediction. Here, F represents a function with fully adaptive thresholds for the feature-item pair $(\boldsymbol{x}, w)$. Specifically, given a query $\boldsymbol{x}$ and item w, we assume a set of predictions $\{\hat{\Delta}^j(\boldsymbol{x}, w)\}_{j=1}^m$ from m different model variants. We then leverage this prediction set to measure uncertainty and establish conformal confidence bounds. For a given query $\boldsymbol{x}$ and item w, let the observed uplift be $\delta(\boldsymbol{x}, w) := (y|a = 1, \boldsymbol{x}, w) - (y|a = 0, \boldsymbol{x}, w)$, and let $\hat{\Delta}(\boldsymbol{x}, w)$ denote the point estimate for the uplift using the mean/median of the predictions in the set or the point prediction from the primary model. Define n as the size of the dataset. We use the sample standard deviation $\hat{\sigma}(\boldsymbol{x}, w)$ from predictions $\{\hat{\Delta}^j(\boldsymbol{x}, w)\}_{j=1}^m$ as the estimated uncertainty. Using $\hat{\sigma}(\boldsymbol{x}, w)$, we can define the nonconformity scores $s(\boldsymbol{x}, w)$ and determine the α-quantile q_α across scores and output $q_\alpha \cdot \hat{\sigma}(\boldsymbol{x}, w)$ as the query-item level threshold with confidence $1 - \alpha$.

We propose two versions of conformal confidence based suppressor, called CoCo-F-α and CoCo-F-β that stand for large number (sample size n) of query-item thresholds. The first one, alpha version, utilizes the observations and model predictions to achieve a more powerful suppressor. The β version handles the case that training data is not accessible for the suppressor and the decisions are made purely based on the model predictions.

CoCo-F-α. In the α version, we use normalized actual residuals as the nonconformity scores and use the quantile from the scores as the threshold. Steps are as follows (i) Compute residuals $r^j(\boldsymbol{x}_i, w_i) = \hat{\Delta}^j(\boldsymbol{x}_i, w_i) - \delta(\boldsymbol{x}_i, w_i)$ and obtain residual set $\{r^j(\boldsymbol{x}_i, w_i)\}_{j=1}^m$. (ii) Compute sample standard deviations $\{\hat{\sigma}(\boldsymbol{x}_i, w_i)\}_{i\in[n]}$ using the residual sets with size m. (iii) Compute normalized residuals as the nonconformity scores: $s(\boldsymbol{x}_i, w_i) = \frac{\hat{\Delta}(\boldsymbol{x}_i, w_i) - \delta(\boldsymbol{x}_i, w_i)}{\hat{\sigma}(\boldsymbol{x}_i, w_i)}$. (iv) Find the sample α-th quantile for scores $\{s(\boldsymbol{x}_i, w_i)\}_{i=1}^{n}$. (v) Output q_α and $\mathrm{T}(\boldsymbol{x}_i, w_i) := q_\alpha \cdot \hat{\sigma}(\boldsymbol{x}_i, w_i),\ \forall i \in [n]$.

CoCo-F-β. In the β version, we use sample deviations which are normalized by standard deviation, as the nonconformity scores and use the quantile from the scores as the threshold. Steps are as follows. (i) Compute sample mean $\{\bar{\Delta}(\boldsymbol{x}_i, w_i)\}_{i \in [n]}$ and sample standard deviations $\{\hat{\sigma}(\boldsymbol{x}_i, w_i)\}_{i \in [n]}$ using the prediction sets with size m. (ii) Compute normalized sample deviations as the nonconformity scores: $s(\boldsymbol{x}_i, w_i) = \frac{\hat{\Delta}(\boldsymbol{x}_i, w_i) - \bar{\Delta}(\boldsymbol{x}_i, w_i)}{\hat{\sigma}(\boldsymbol{x}_i, w_i)}$. (iii) Find the sample α-th quantile for scores $\{s(\boldsymbol{x}_i, w_i)\}_{i=1}^{n}$. (iv) Output quantile q_α and $\mathrm{T}(\boldsymbol{x}_i, w_i) := q_\alpha \cdot \hat{\sigma}(\boldsymbol{x}_i, w_i), \forall i \in [n]$.

3 Evaluation

3.1 Offline Evaluation

Data Simulation: To evaluate the performance of CoCo suppressor methods, we simulate synthetic data with various configurations. Each dataset consists of $N = 1000$ samples, with predicted and observed values generated under specified noise levels for prediction (σ^2_{pred}) and observation (σ^2_{obs}). Additional biases are incorporated to assess robustness. In our simulation, we experiment with $\sigma^2_{\text{pred}} \in \{0.1, 1.0, 10.0, 50.0\}$ and fix $\sigma^2_{\text{obs}} = 0.01$. We also study model prediction biases controlled by $\{0\%, 1\%, 5\%, 10\%, 20\%\}$, representing the percentage change relative to the true means.

Evaluation Metrics: We evaluate the methods using two metrics: (i) **Positive Predictive Value (PPV)**: the precision of suppression, calculated as the proportion of correctly suppressed samples among all suppressed samples; (ii) **True Positive Rate (TPR)**: the recall of suppression, calculated as the proportion of correctly suppressed negative samples among all true negatives. Metrics are averaged over 100 repetitions to ensure statistical robustness.

Performance Summary: Tables 1 and 2 present the summary results for PPV and TPR. All suppressors are valid across all settings, outperforming random guessing (50% PPV), with precision exceeding 90% in favorable settings (low prediction variance and bias). However, suppressors tend to be conservative—achieving high PPV but low TPR. In general, α-suppressors exhibit stronger suppression capability, while β-suppressors are more conservative. Notably, TPR is sensitive to prediction uncertainty, decreasing substantially (below 20%) when prediction variance exceeds 1.0.

The results indicate that CoCo-1 achieves the best performance when σ^2_{pred} is moderate and homogeneous. CoCo-M suppressors remain competitive even when the context variable does not represent the true underlying groups. Our simulations do not demonstrate the superiority of CoCo-F suppressors; under Gaussian distributions with homogeneous variance, CoCo-F offers no additional benefit. Further research is needed to evaluate CoCo-F's potential advantages. Given the trade-off between performance and implementation complexity in production environments (e.g., training a set of models in CoCo-F to estimate model uncertainty), we proceed with CoCo-M for online experimentation. In practice, we recommend CoCo-1 for homogeneous settings, CoCo-M when known contextual groups exist (e.g., regions on the SRP), and CoCo-F when model ensembles are available and variance is heterogeneous across query-item pairs.

Config		CoCo-1		CoCo-M		CoCo-F	
σ^2	Bias	α	β	α	β	α	β
0.1	0%	99.36±0.07	**100.00±0.00**	99.39±0.07	**100.00±0.00**	99.37±0.06	**100.00±0.00**
0.1	1%	99.32±0.07	**100.00±0.00**	99.35±0.07	**100.00±0.00**	99.34±0.07	**100.00±0.00**
0.1	5%	99.17±0.08	**100.00±0.00**	99.14±0.08	**100.00±0.00**	99.18±0.08	**100.00±0.00**
0.1	10%	99.26±0.07	**100.00±0.00**	99.28±0.07	**100.00±0.00**	99.27±0.07	**100.00±0.00**
0.1	20%	99.21±0.08	**100.00±0.00**	99.24±0.08	**100.00±0.00**	99.16±0.09	**100.00±0.00**
1.0	0%	94.14±0.32	**98.03±0.24**	93.90±0.31	97.73±0.24	93.86±0.29	97.76±0.25
1.0	1%	93.94±0.32	97.71±0.34	93.71±0.30	97.61±0.34	93.75±0.32	**97.85±0.33**
1.0	5%	93.57±0.34	**97.67±0.31**	93.67±0.35	97.27±0.29	94.05±0.34	97.31±0.36
1.0	10%	93.58±0.27	97.19±0.31	93.42±0.26	**97.35±0.33**	93.49±0.26	96.97±0.34
1.0	20%	93.90±0.36	**97.71±0.25**	93.69±0.38	97.65±0.28	93.72±0.37	97.45±0.29
10	0%	73.72±0.89	**75.00±0.88**	73.82±0.92	74.90±0.92	73.69±0.91	74.55±0.92
10	1%	74.20±0.82	**75.61±0.91**	73.99±0.85	75.15±0.89	74.25±0.80	75.28±0.88
10	5%	74.88±0.86	75.39±0.85	74.53±0.87	**75.76±0.87**	74.98±0.79	75.63±0.89
10	10%	73.52±0.98	73.78±1.03	73.55±0.98	**74.06±1.00**	72.95±1.01	73.84±1.06
10	20%	75.16±0.93	75.73±1.06	74.74±0.95	75.44±0.98	75.31±0.98	**75.84±1.00**
50	0%	61.10±0.90	61.19±0.93	60.90±0.95	**61.42±0.92**	60.87±0.87	60.56±0.88
50	1%	59.96±0.94	60.59±0.93	60.16±0.99	60.18±0.93	60.46±0.85	**60.62±0.97**
50	5%	**63.53±1.04**	63.42±1.06	62.96±1.15	62.78±1.11	62.49±1.03	62.57±1.10
50	10%	61.53±0.76	61.37±0.72	61.55±0.80	61.40±0.77	**62.00±0.78**	61.84±0.84
50	20%	59.91±0.89	59.87±0.96	59.69±0.90	59.75±0.99	60.20±0.99	**60.73±1.01**

Table 1: Performance Summary Table for PPV

Config		CoCo-1		CoCo-M		CoCo-F	
σ^2	Bias	α	β	α	β	α	β
0.1	0%	**60.11±0.36**	10.07±0.24	**60.11±0.38**	10.14±0.25	59.98±0.37	10.25±0.27
0.1	1%	**59.97±0.48**	10.42±0.23	59.93±0.47	10.48±0.22	59.65±0.45	10.32±0.21
0.1	5%	60.30±0.33	9.79±0.22	**60.37±0.32**	10.02±0.22	60.12±0.33	9.97±0.26
0.1	10%	59.17±0.35	10.06±0.24	**59.31±0.36**	10.14±0.24	58.98±0.37	10.24±0.26
0.1	20%	**59.00±0.32**	10.18±0.25	58.86±0.32	10.38±0.26	58.74±0.32	10.38±0.28
1.0	0%	22.79±0.34	9.79±0.23	**22.98±0.37**	9.83±0.25	22.72±0.34	9.75±0.25
1.0	1%	23.08±0.33	9.71±0.23	**23.25±0.31**	9.82±0.25	22.95±0.30	10.00±0.24
1.0	5%	**23.13±0.37**	10.15±0.24	23.12±0.38	10.28±0.23	22.87±0.38	10.03±0.22
1.0	10%	23.37±0.30	9.99±0.20	**23.39±0.33**	10.21±0.19	23.12±0.32	10.01±0.22
1.0	20%	22.67±0.26	9.65±0.22	**22.72±0.25**	9.79±0.23	22.59±0.24	9.91±0.23
10	0%	8.57±0.21	7.32±0.21	**8.67±0.21**	7.43±0.19	8.43±0.20	7.19±0.19
10	1%	8.95±0.26	7.59±0.21	**9.04±0.24**	7.69±0.21	8.91±0.25	7.62±0.26
10	5%	9.04±0.25	7.47±0.21	**9.16±0.25**	7.61±0.19	9.01±0.24	7.59±0.22
10	10%	9.08±0.28	7.77±0.21	**9.27±0.27**	8.12±0.22	9.08±0.27	7.81±0.25
10	20%	8.72±0.25	7.55±0.21	**8.80±0.25**	7.60±0.21	8.63±0.24	7.41±0.22
50	0%	6.67±0.19	6.33±0.16	**6.72±0.19**	6.44±0.17	6.67±0.18	6.30±0.17
50	1%	6.10±0.20	5.91±0.19	**6.23±0.21**	6.18±0.20	6.17±0.19	5.93±0.18
50	5%	**6.52±0.20**	6.33±0.19	6.50±0.21	6.32±0.19	6.39±0.21	6.09±0.20
50	10%	6.59±0.21	6.40±0.19	6.70±0.21	6.56±0.18	**6.71±0.21**	6.22±0.18
50	20%	6.40±0.14	6.20±0.16	**6.50±0.14**	6.20±0.16	6.34±0.16	5.99±0.15

Table 2: Performance Summary Table for TPR

3.2 Online Evaluation

Workflow Setup: We implement the proposed methodology in a production system (Figure 2). Our data collection involves both factual and counterfactual observations across different search contexts. Factual data are obtained from actual item impressions, comprising 95% of the traffic. For counterfactual data, we reserve 5% exploration traffic for random shuffling, allowing us to observe outcomes when an item that would have been shown is withheld due to random reordering. We note that the counterfactual model $f_{a=0}$, trained on a smaller exploration dataset, may exhibit higher prediction variance than the factual model $f_{a=1}$. The CoCo thresholding mechanism naturally accounts for this, as higher prediction noise in either model propagates into larger nonconformity scores, resulting in more conservative suppression decisions. This yields contextual request data paired with rewards for both shown ($a = 1$) and withheld ($a = 0$) items. For causal effect modeling, we train two separate deep learning models on the factual and counterfactual datasets. Our model uses 12 categorical, 10 numeric, and transformer-encoded query features. The architecture is a 2-layer MLP (256-dim) with 3 multi-task heads, trained via Adam (lr=0.001) on 14 days of data. During online inference, these models predict the estimated item value. Suppression thresholds are generated using CoCo-M, with region-specific thresholds for different areas of the SRP to account for varying engagement patterns. The computational overhead is offset by significant business gains (Table 3).

A/B Test Design: *Control*: we employ a heuristic suppression method where the baseline for widget uplift calculation is derived from the average reward across all widgets within certain dimensions (denoted as y_{avg}), such as country. Suppression decisions are made when $f_{a=1}(\boldsymbol{x}, w) < y_{\text{avg}}$. *Treatment 1 (T1)*: we apply our uplift estimation but with heuristic 0 as the suppression threshold. *Treatment 2 (T2)*: we apply CoCo-M as in Section 2.3.2. The setup is the same for Control and Treatments except for the suppression method. Traffic was randomly split at the user level across C/T1/T2.

A/B Test Results: Both T1 and T2 outperform the control, and T2 further outperforms T1 across metrics, demonstrating the benefit of incorporating model uncertainty. In Table 3, the 4-week global online test shows statistically significant improvements in business metrics (e.g., increased revenue and profit) and customer engagement (e.g., reduced search abandonment). While the relative improvements appear modest in percentage, at the scale of a major e-commerce platform serving hundreds of millions of users, even small percentage gains translate to substantial business impact. T1-C and T2-C denote lift over control; PPR measures the Bayesian probability that the treatment effect is positive; PPR > 66.6% or < 33.3% indicates statistical significance under a one-sided test.

Metric	T1-C (%)	PPR (%)	T2-C (%)	PPR (%)
Revenue	0.001%	48.4%	0.03%	89.2%
Long-term Revenue	0.006%	64.3%	0.025%	89.9%
Purchased Units	0.009%	71.3%	0.022%	77.7%
Ads Revenue	0.018%	71.9%	0.037%	88.0%
Short-term Profit	0%	50.2%	0.016%	58.5%
Long-term Profit	0.005%	73.7%	0.017%	71.5%
Search Abandonment Rate	-0.02%	8.11%	-0.01%	27.2%

Table 3: Metrics Analysis (PPR > 66.6% or PPR < 33.3% is Statistically Significant)

4 Discussion

4.1 Unified Ranking and Suppression

The CoCo framework, introduced for content suppression, naturally extends to a unified ranking and suppression system. In the above suppression-only setup (Section 2), we train separate factual and counterfactual models and suppress items when the estimated uplift $\hat{\Delta}(\boldsymbol{x}, w) = f_{a=1}(\boldsymbol{x}, w) - f_{a=0}(\boldsymbol{x}, w)$ falls below a CoCo threshold. A key observation is that the factual model $f_{a=1}(\boldsymbol{x}, w)$, which estimates the expected reward of *showing* an item, can directly serve as a ranking score, i.e., items with higher estimated value of being shown should be ranked higher. This leads to a unified scoring function for item w:

$$\text{score}(w) = f_{a=1}(\boldsymbol{x}, w) \cdot \mathbb{1}\left[\hat{\Delta}(\boldsymbol{x}, w) \geq \mathrm{T}(\boldsymbol{x}, w)\right], \qquad (2)$$

where ranking is determined by the factual model and suppression is applied via the CoCo threshold on the uplift estimate. One can view this as a two-stage constrained optimization: first ranking by factual value, then suppressing items whose incremental value (factual value minus counterfactual value) does not meet the confidence threshold. This unified scoring function decomposes the problem along its natural axes. Ranking is a decision conditional on display (the $a = 1$ regime), so the factual model $f_{a=1}(\boldsymbol{x}, w)$ is the appropriate ranking score, while the uplift $\hat{\Delta}(\boldsymbol{x}, w)$ is the appropriate object for the show-vs.-withhold decision. Ranking by $\hat{\Delta}$ would conflate the two regimes, yielding orderings that do not preserve the ranking of realized user experience under display. Driving both decisions from the same pair of causal effect models also guarantees consistency.

4.2 Online Evaluation of Unified System

A/B Test Design: The control uses the suppression-only setup described in Section 3.2, with separate models for ranking and suppression. We initially tested 16 treatments via adaptive experimentation to isolate the effect of each change dimension (e.g., objective function, features, suppression threshold method), and selected the best-performing treatment for the final 4-week global validation (Table 4). The treatment unifies ranking and suppression using the factual model for ranking and both factual and counterfactual models for suppression, with CoCo thresholds automatically refreshed daily per slot to account for data drift. The treatment also adopts a unified objective function incorporating view-through rate (VTR) as an engagement signal and adds extra 3C features, i.e., Contextual features (e.g., page number), Content features (e.g., content ontology), and Customer features (e.g., recent search queries).

	Control	Treatment
Ranking model	Separate	Factual
Suppression model	Separate	Factual/Counterfactual
Objective	Reward	Reward + $w\cdot$ VTR
Features	Base	Base + extra 3C
Suppression threshold	Static	CoCo per-slot (auto)

Table 4: Unified Ranking and Suppression A/B Test Setup

A/B Test Results: In Table 5, the unified system achieves statistically significant improvements in revenue (PPR = 90.2%), long-term revenue (PPR = 90.0%), advertisement revenue (PPR = 100.0%), and short-term profit (PPR = 91.2%), while maintaining flat purchased units and long-term profit. Additionally, the unified modeling reduces the number of trained models by over 70% and offline pipeline costs by over 40%. These results demonstrate that the CoCo framework provides both business and operational value when extended from suppression to joint ranking and suppression.

Metric	T-C (%)	Prob of Positive Return (PPR)
Revenue	0.031%	90.2%
Long-term Revenue	0.02%	90.0%
Purchased Units	-0.005%	37.9%
Ads Revenue	0.094%	100.0%
Short-term Profit	0.029%	91.2%
Long-term Profit	0.005%	63.3%
Search Click Rate	0.007%	85.4%

Table 5: Unified system (Treatment) vs. separate ranking/suppression models (Control)

5 Conclusion

In this paper, we propose CoCo, a content suppression mechanism that incorporates prediction uncertainty into uplift-based suppression decisions via conformal prediction. Our approach derives distribution-free thresholds at varying contextual granularity, mitigating over-suppression without requiring auxiliary uncertainty models. Both offline and online evaluations confirm statistically significant gains in revenue and customer engagement. We further show that the framework extends to unified ranking and suppression, achieving additional business and operational gains in production.

References

[1] Anastasios N Angelopoulos and Stephen Bates. 2021. A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *arXiv preprint arXiv:2107.07511* (2021).

[2] Stephen Bonner and Flavian Vasile. 2018. Causal embeddings for recommendation. In *Proceedings of the 12th ACM conference on recommender systems*. 104–112.

[3] Jiaju Chen, Wang Wenjie, Chongming Gao, Peng Wu, Jianxiong Wei, and Qingsong Hua. 2024. Treatment Effect Estimation for User Interest Exploration on Recommender Systems. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1861–1871.

[4] Pierre Gutierrez and Jean-Yves Gérardy. 2017. Causal inference and uplift modelling: A review of the literature. In *International conference on predictive applications and APIs*. PMLR, 1–13.

[5] Sören R Künzel, Jasjeet S Sekhon, Peter J Bickel, and Bin Yu. 2019. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the national academy of sciences* 116, 10 (2019), 4156–4165.

[6] Zhou Qin, Kai Yuan, Pratik Lahiri, and Wenyang Liu. 2024. Cooperative Multi-Agent Deep Reinforcement Learning in Content Ranking Optimization. *arXiv preprint arXiv:2408.04251* (2024).

[7] Weijia Zhang, Jiuyong Li, and Lin Liu. 2021. A unified survey of treatment effect heterogeneity modelling and uplift modelling. *ACM Computing Surveys (CSUR)* 54, 8 (2021), 1–36.