RoLGaL: Robust Learning for CVR Prediction via Guardrail-Guided Labeling

Xiaoli Gao, Qiuping Xu, Amir Roshankar, Rong Jin
{emmagao,xqiuping,aroshank,rongjinml}@meta.com
Meta
Bellevue, WA, USA

Abstract

Modern ads ranking systems update models daily or hourly to track shifting user behavior. However, continual training exposes CVR prediction to heterogeneous and non-stationary label noise from accidental clicks, delayed attribution, and system artifacts. Such noisy updates can overwrite stable user preference patterns, leading to catastrophic forgetting.

We introduce **RoLGaL** (**Ro**bust **L**earning with **Ga**uardrail-guided **L**abeling), a framework for robust continual CVR training. RoLGaL combines label diffusion, guardrail-guided correction, and a multi-head auxiliary architecture to construct adjusted labels for robust auxiliary training. Rather than replacing the serving target, RoLGaL uses these adjusted labels only in an auxiliary task, improving representation robustness while preserving primary-head calibration and adding zero inference latency.

On the Criteo Attribution dataset, RoLGaL improves AUC by 0.25% and reduces LogLoss by 0.55%. Under simulated label corruption, it achieves 0.31% and 0.34% AUC lifts at 1% and 5% noise, respectively. In large-scale A/B testing at a leading advertising platform, RoLGaL achieves a statistically significant 0.23% lift in a primary revenue-related business metric with zero inference overhead. It has since been deployed across 10+ production CVR models, demonstrating practical impact at scale.

Keywords

Online Advertising, Continual Learning, Catastrophic Forgetting, Label Noise, Bayesian De-Noising, Recommender Systems

ACM Reference Format:
Xiaoli Gao, Qiuping Xu, Amir Roshankar, Rong Jin. 2026. RoLGaL: Robust Learning for CVR Prediction via Guardrail-Guided Labeling. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '26), August 9–13, 2026, Jeju, South Korea.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/XXXXXX.XXXXXX

1 Introduction

Post-click conversion (CVR) models are central to online advertising, estimating the likelihood of valuable user actions such as purchases or app installs. To adapt to shifting user behavior and market dynamics, production systems continually refresh these models on recent data, often daily or hourly.

However, continual training also exposes CVR models to heterogeneous and non-stationary label noise. Observed conversion labels may be affected by accidental clicks, delayed or cross-device attribution, sparse feedback, and system-level logging artifacts. Because this noise varies across time, users, devices, and attribution windows, models must learn from recent data without allowing corrupted supervision to dominate the update. When noisy labels enter the training stream, their gradients can overwrite stable historical patterns, leading to noise overfitting and catastrophic forgetting.

Existing approaches typically address noisy-label learning or continual learning separately. Noisy-label methods often assume stationary corruption or rely on clean validation signals, while continual-learning methods usually assume reliable incoming data. In large-scale ads ranking, both assumptions are restrictive: labels are delayed and non-stationary, clean validation signals are difficult to obtain, and preserving calibrated CVR estimates is critical for downstream auction bidding.

We propose RoLGaL (Robust Learning with Guardrail-Guided Labeling), a calibration-preserving framework for robust continual CVR training. RoLGaL combines three components: (1) **label diffusion**, a diffusion-inspired probabilistic correction step that produces denoised soft targets; (2) **temporal guardrails**, which compare recovered labels against stable reference signals to flag high-risk updates; and (3) a **multi-head auxiliary architecture**, which aggregates multiple denoising views to improve robustness under heterogeneous noise. RoLGaL is implemented as an auxiliary task in a multi-task learning setup: the primary head trains on observed labels to preserve calibration, while the auxiliary head learns from adjusted labels to regularize the shared representation. This design places robustness in the representation learning process while leaving the bid-critical serving target unchanged. At inference time, the auxiliary head is discarded, yielding zero latency overhead.

Our contributions are:

- We introduce RoLGaL, a calibration-preserving robust continual training framework that combines probabilistic label correction and guardrail-guided adjustment for noisy CVR supervision.
- Offline experiments on Criteo show a 0.25% AUC improvement and stronger robustness under simulated label corruption, with 0.31% and 0.34% AUC lifts at 1% and 5% noise.
- A large-scale production A/B test shows a statistically significant 0.23% lift in a primary revenue-related business metric with zero inference latency overhead, followed by deployment across 10+ production CVR models.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
AdKDD '26, Jeju, South Korea
© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-XXXX-X/26/08
https://doi.org/10.1145/XXXXXX.XXXXXX

2 Related Work

Existing methods in conversion prediction typically treat noisy-label learning, continual learning, and knowledge distillation in isolation. In learning from noisy labels, approaches like transition matrices [14] or loss-based sample selection [5, 10] often assume stationary noise patterns. Recent recommendation models [2, 6] improve implicit feedback modeling but do not directly address non-stationary supervision in continual ads training. This is limiting in ads systems, where noise can fluctuate by time, device, attribution window, and logging condition. RoLGaL addresses this dynamic environment through instance-level soft-label correction and temporal guardrails.

Continual learning methods use regularization [9] or replay [12] to mitigate catastrophic forgetting. However, they usually assume reliable incoming data; replaying noisy samples may preserve corrupted patterns. Similarly, meta-learning approaches for sample reweighting [15, 16] often require clean validation sets, which are difficult to obtain in streaming ads systems. RoLGaL instead validates incoming supervision against temporal reference signals and applies corrected targets through an auxiliary task, keeping the primary task trained on observed labels for calibration.

RoLGaL is also related to knowledge distillation [4, 7, 19] and label smoothing [17]. Unlike standard distillation, which transfers teacher supervision broadly, RoLGaL uses dynamic checkpoint-based references as guardrails and applies correction selectively to high-risk samples. Unlike label smoothing, which applies a uniform penalty to all labels, RoLGaL produces instance-dependent adjusted labels. We compare against both methods experimentally. The main distinction of RoLGaL is its production-oriented integration: guardrail-guided label adjustment regularizes the shared representation while keeping the serving head and inference path unchanged.

3 Problem Formulation

In continual ads ranking, time is discretized into steps $t = 1, \ldots, T$. At each step, the system receives a batch $\mathcal{D}_t = \{(x_i, y_i)\}_{i=1}^{N_t}$, where x_i denotes user-ad features and $y_i \in \{0, 1\}$ is the observed conversion label. The model $f_\theta(x) \in [0, 1]$ is initialized from the previous checkpoint and updated on the current batch using binary cross-entropy.

Observed labels may differ from latent user intent $\tilde{y}_i$ due to accidental clicks, delayed or cross-device attribution, logging artifacts, and other system-level effects. We model this as instance- and time-dependent noise:

$$P(y_i \neq \tilde{y}_i \mid x_i) = \eta(x_i, t).$$

Unlike standard noisy-label settings with stationary corruption, ads label noise changes over time and across segments. Continual training on such streams can cause both noise overfitting and catastrophic forgetting, where gradients from recent noisy batches overwrite stable historical preference patterns.

Our goal is to minimize the expected risk with respect to latent intent,

$$\mathcal{J}(\theta) = \mathbb{E}_{(x,\tilde{y})} \left[\ell(f_\theta(x), \tilde{y}) \right],$$

while only observing corrupted labels. In production CVR systems, however, directly replacing observed labels may shift the predicted probability distribution and harm calibration. RoLGaL therefore approximates this objective by constructing denoised surrogate labels for auxiliary regularization, while keeping the primary prediction task trained on observed labels.

4 RoLGaL Framework

To address noisy continual training and catastrophic forgetting, we introduce **RoLGaL**. RoLGaL constructs adjusted labels for robust auxiliary training rather than directly replacing the primary training target. The framework refines raw observed labels through three sequential stages.

(1) **Label Diffusion:** A diffusion-inspired probabilistic correction step that produces denoised soft labels.
(2) **Guardrail-guided Judgment:** Validating the recovered labels against stable reference signals.
(3) **Adaptive Label Adjustment:** Combining the recovered signal and guardrail judgment into a final adjusted label y^a for the auxiliary task.

To improve robustness against heterogeneous noise, we extend this pipeline into a **Multi-Head architecture**. Instead of relying on a single set of correction parameters, multiple parallel heads run the above stages with different noise and guardrail sensitivities. This provides multiple denoising views while keeping the serving path unchanged.

Finally, we integrate these components using a **Joint Training Strategy**. The adjusted labels supervise an auxiliary task, while the primary ranking task continues to use raw observed labels. This design allows RoLGaL to improve representation robustness without directly shifting the primary prediction target, helping preserve calibration for production CVR serving. Figure 1 illustrates the overall framework.

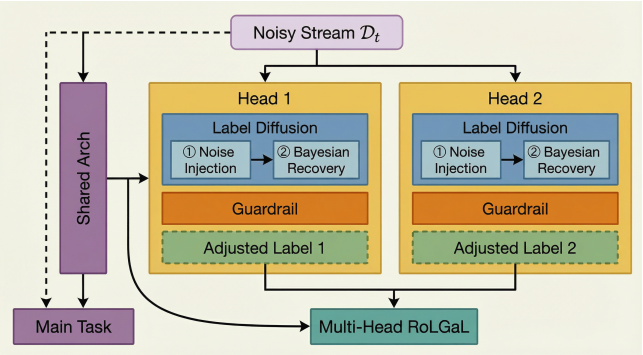

Figure 1: RoLGaL Framework. The framework operates within a shared-bottom multi-task learning architecture. The noisy stream is processed by parallel RoLGaL heads, shown here with two heads for illustration. Each head applies label diffusion and guardrail-guided judgment to produce adjusted labels. The main task is supervised by raw observed labels y^{obs} to preserve calibration, while the RoLGaL auxiliary task is supervised by adjusted labels y^a to improve representation robustness. At serving time, the auxiliary head is discarded, yielding zero inference overhead.

4.1 Label Diffusion

The Label Diffusion module produces soft corrected labels from noisy observations. Inspired by diffusion-style denoising, it uses a lightweight one-step probabilistic relaxation: an observation model that accounts for label noise, followed by Bayesian recovery of a soft target. Unlike full generative diffusion models, our goal is not to learn an iterative reverse process, but to obtain an efficient denoised label estimate suitable for continual ranking training.

Step 1: Noise Injection (Observation Model). For each latent conversion intent $\tilde{y}_i \in \{0, 1\}$, we view the observed label y_i^{obs} as a noisy relaxed observation generated around the latent label:

$$y_i^{obs} \mid \tilde{y}_i \sim N(\tilde{y}_i, \sigma^2), \tag{1}$$

where σ^2 controls the assumed label-noise level.

Step 2: Bayesian Recovery. Given the noisy observation, we compute the posterior probability that the latent label is positive:

$$y_i^* = P(\tilde{y}_i = 1 \mid y_i^{obs}) = \frac{1}{1 + \left(\frac{1-p_i}{p_i}\right) \exp\left(\frac{1-2y_i^{obs}}{2\sigma^2}\right)}, \tag{2}$$

where p_i denotes the prior conversion probability, approximating $E[\tilde{y}_i \mid x_i]$ for input feature x_i. The resulting y_i^* is a soft corrected label that smooths unreliable binary supervision before guardrail-based validation.

4.2 Guardrail-guided Judgment

While Label Diffusion provides a soft corrected label, it cannot by itself determine whether the correction is consistent with historically learned user-ad patterns. To address this, we introduce a **guardrail label** y_i^g, a temporal reference signal used to detect high-risk deviations. This signal can be derived from the **production model** or previous checkpoint predictions, a **teacher model** that provides soft supervision, or **heuristic rules** that encode domain constraints such as dwell-time thresholds.

In our framework, y_i^g acts as a conservative reference rather than a clean-label oracle. Unlike model ensembles that mainly reduce variance, or self-training methods that may amplify memorized noise, the guardrail compares the recovered label against a slower-moving reference signal. This helps identify samples whose corrected labels strongly contradict established patterns, while still allowing moderate changes that may reflect genuine distribution shift.

Specifically, we evaluate the quality of the recovered label y_i^* by computing its distance from the guardrail. If this distance exceeds a stability threshold τ, the sample contradicts established patterns and is flagged as **Risky** ($r_i = 1$); otherwise, it is deemed **Safe** ($r_i = 0$):

$$r_i = I\left(\text{Dist}(y_i^*, y_i^g) > \tau\right). \tag{3}$$

This distance metric can be an absolute difference or a log odds ratio. In our experiments, we use the log odds ratio and set $\tau \in [1, 2]$. This selective judgment reduces the impact of abrupt noise-driven updates while allowing the model to adapt when the recovered signal remains consistent with the temporal reference.

4.3 Adaptive Label Adjustment

The final stage combines the recovered label y_i^* and the guardrail-guided risk indicator r_i into an **adjusted label** y_i^a. This label is used as the supervision target for the RoLGaL auxiliary task.

We use an adaptive mixing strategy:

$$y_i^a = (1 - r_i) \cdot y_i^* + r_i \cdot \left(c_i y_i^g + (1 - c_i) y_i^*\right), \tag{4}$$

where $c_i \in [0, 1]$ controls how strongly the guardrail label is trusted for risky samples. Safe samples ($r_i = 0$) use the recovered label y_i^*, while risky samples ($r_i = 1$) interpolate between the recovered label and the guardrail label.

In our standard setup, we set $c_i = 1$ for confirmed high-risk samples, simplifying the logic to a direct gating mechanism:

$$y_i^a = \begin{cases} y_i^* & \text{if } r_i = 0 \text{ (Safe)} \\ y_i^g & \text{if } r_i = 1 \text{ (Risky)} \end{cases} \tag{5}$$

The adjusted label is applied only to the auxiliary task, allowing RoLGaL to regularize the shared representation without directly changing the primary calibrated training target.

4.4 Multi-Head Architecture

A single noise variance σ^2 or guardrail threshold τ may not fit all samples in a non-stationary ads stream. To reduce sensitivity to any single correction setting, RoLGaL uses a lightweight multi-head auxiliary design. Each head $k \in \{1, \ldots, K\}$ runs the same correction pipeline with a distinct configuration:

$$\Phi_k = \{\sigma_k^2, \tau_k\}. \tag{6}$$

This produces head-specific adjusted labels $y_{i,k}^a$, each corresponding to a different noise and guardrail sensitivity.

The auxiliary loss aggregates the supervision from all heads:

$$\mathcal{L}_{aux}(\mathcal{B}) = \sum_{k=1}^{K} \omega_k \cdot \ell_{BCE}(\hat{p}^a; \mathcal{B}_k), \tag{7}$$

where ω_k is the head weight, typically uniform with $\omega_k = 1/K$. For each head,

$$\ell_{BCE}(\hat{p}^a; \mathcal{B}_k) = -\sum_{i \in \mathcal{B}} \left[y_{i,k}^a \log \hat{p}_i^a + (1 - y_{i,k}^a) \log(1 - \hat{p}_i^a) \right]. \tag{8}$$

This averaged auxiliary objective prevents the representation from depending on a single correction view. In practice, we keep K small to avoid unnecessary complexity; unless otherwise specified, we use $K = 2$.

4.5 Implementation: Joint Training Strategy

Directly replacing observed labels with adjusted labels may shift the predicted distribution, which is undesirable in ads systems where auction bidding depends on calibrated CVR estimates. To balance robustness and calibration, RoLGaL is implemented as an auxiliary task in a Shared-Bottom multi-task learning (MTL) architecture.

Architecture & Joint Loss. A shared encoder f_θ maps input x_i to a latent representation h_i. This feeds into two branches: a **Primary Head** predicting $\hat{p}^m = \text{Sigmoid}(h_i^T w_{main})$ trained on raw labels y_i^{obs} to preserve the serving target distribution, and a **RoLGaL Auxiliary Head** predicting $\hat{p}^a = \text{Sigmoid}(h_i^T w_{aux})$

trained on adjusted labels y_i^a to regularize the shared representation. The model is trained end-to-end via:

$$\mathcal{L}_{total} = \ell_{BCE}(\hat{p}^m; \mathcal{B}) + \lambda \mathcal{L}_{aux}(\mathcal{B}) \tag{9}$$

where λ controls the strength of the auxiliary robustness objective. By jointly optimizing these objectives, RoLGaL improves the shared representation using corrected supervision while keeping the production-serving head trained on empirical labels.

Inference. At serving time, the auxiliary head is discarded, yielding zero inference latency overhead.

5 Experiments

We evaluate RoLGaL in two complementary settings. First, we conduct public benchmark experiments on the Criteo Attribution Dataset to evaluate offline ranking quality, robustness to injected label noise, and component-level contributions. Second, we validate RoLGaL in a large-scale production ads system through online A/B testing and subsequent deployment across multiple conversion ranking models.

This design separates controlled methodological validation from deployment validation. The public benchmark experiments use reproducible noisy-label stress tests to compare methods as supervision quality degrades. These synthetic corruptions do not capture every form of production attribution noise, but they provide a controlled setting for measuring robustness. The production evaluation then tests whether the same auxiliary-task design translates into practical gains under real ads-system constraints, including calibration requirements, serving latency, and online business impact. For public experiments, we report AUC and LogLoss. For production validation, we report online lift on a primary revenue-related business metric, together with deployment efficiency and rollout evidence.

5.1 Public Benchmark Setup

Criteo Attribution Dataset. We conduct our primary controlled experiments on the Criteo Attribution Dataset [3], a widely adopted benchmark for industrial conversion prediction. The dataset contains over 16.5M chronologically ordered interactions with sparse conversion labels, making it suitable for streaming CVR prediction under label noise.

Baseline Model. We use DCN-v2 [18], a strong backbone for conversion prediction, as the base architecture for all methods. This ensures that observed gains come from RoLGaL's robust-training strategy rather than backbone differences.

Streaming Setup and Guardrail Generation. We simulate continual learning using a chronological split. The first 14 of 31 days are used as a warm-up period to train an initial stable model. The next 7 days are used to train the baseline and RoLGaL models from the warm-up checkpoint. For the remaining 10 days, we use streaming evaluation: the model is updated on daily batch $\mathcal{D}_t$ and evaluated on the subsequent day $\mathcal{D}_{t+1}$.

To generate guardrail labels without data leakage, we use historical self-supervision. For each incoming batch $\mathcal{D}_t$, the frozen checkpoint from the previous step, $f_{\theta_{t-1}}$, performs forward inference before any update on the current batch. These predictions serve as guardrail labels $\{y_i^g\}$, anchoring the judgment to the model's previously learned distribution.

Noise Injection Simulation. To evaluate robustness under controlled corruption, we flip each training label with probability η_{noise}:

$$P(y^{obs} \neq \tilde{y}) = \eta_{noise}.$$

We evaluate $\eta_{noise} \in \{0.01, 0.05, 0.10\}$ to test model stability from mild to aggressive label corruption.

5.2 Public Benchmark Results

We compare RoLGaL with the baseline model and two common robust-training baselines: Label Smoothing (LS) and Knowledge Distillation (KD). Each configuration is evaluated across three independent trials; we report mean performance, standard deviation, and relative lift. Unless otherwise specified, Multihead RoLGaL uses $K = 2$ heads.

Table 1: Main Results on Criteo. We report Mean ± Std and relative lift from the baseline across 3 trials.

Method	AUC (↑)	LogLoss (↓)
Baseline	0.9027 ± 0.0002	0.12563 ± 0.00013
Label Smoothing	0.9027 ± 0.0003 (+0.00%)	0.12556 ± 0.00008 (-0.08%)
Knowledge Distillation	0.9047 ± 0.0003 (+0.22%)	**0.12488 ± 0.00025 (-0.60%)**
Singlehead RoLGaL	0.9048 ± 0.0003 (+0.23%)	0.12506 ± 0.00017 (-0.46%)
Multihead RoLGaL	**0.9050 ± 0.0003 (+0.25%)**	0.12494 ± 0.00021 (-0.55%)

Main Performance. Table 1 shows that LS provides limited benefit on clean Criteo data, while KD is a strong baseline, approaching RoLGaL in AUC and achieving the best point-wise LogLoss. Multihead RoLGaL obtains the best AUC lift (+0.25%), and Singlehead RoLGaL is comparable to KD (+0.23% vs. +0.22%). These results show that RoLGaL is competitive with strong teacher-based supervision under clean benchmark conditions. We next evaluate robustness under injected label noise, where selective guardrail-guided correction becomes more important.

5.3 Robustness to Label Noise

To evaluate robustness under controlled label corruption, we inject random label noise into the Criteo training stream. For each training sample, we flip the observed label with probability η, while keeping the evaluation labels unchanged. This setup is a controlled stress test rather than a complete model of production attribution noise.

Table 2: Robustness under Injected Noise ($\eta = 0.01$). RoLGaL outperforms LS and KD under noisy supervision.

Method	AUC (↑)	Lift
Baseline	0.8989 ± 0.0002	–
Label Smoothing	0.8988 ± 0.0001	-0.01%
Knowledge Distillation	0.9006 ± 0.0002	+0.19%
Multihead RoLGaL	**0.9016 ± 0.0001**	**+0.31%**

Comparison with LS and KD. Table 2 shows that LS fails to improve over the noisy baseline, while KD achieves a +0.19% AUC lift. In contrast, Multihead RoLGaL achieves a +0.31% lift, a 63% relative improvement over KD. This suggests that selective label correction is more robust than uniform smoothing or broad teacher supervision when the training stream contains corrupted labels.

KD remains highly competitive on relatively clean benchmark data, where the teacher provides a useful stabilizing signal. Under explicit label corruption, however, teacher supervision is less targeted because it does not directly identify which samples should be corrected. RoLGaL applies correction selectively: label diffusion produces a soft corrected target, and the guardrail flags high-risk samples whose recovered labels strongly deviate from temporal references.

Noise-Level Sensitivity. We further vary the label-flipping probability $\eta \in \{0.01, 0.05, 0.10\}$ and compare Singlehead and Multihead RoLGaL against the corresponding noisy baseline.

Table 3: Robustness under Injected Label Noise.

Noise (η)	Baseline	Singlehead RoLGaL ($K = 1$)	Multihead RoLGaL ($K = 2$)
0.01	0.8989 ± 0.0002	0.9013 ± 0.0002 (+0.27%)	**0.9016 ± 0.0001 (+0.31%)**
0.05	0.8900 ± 0.0002	0.8930 ± 0.0002 (+0.34%)	**0.8931 ± 0.0001 (+0.34%)**
0.10	0.8814 ± 0.0003	0.8850 ± 0.0003 (+0.41%)	**0.8852 ± 0.0004 (+0.43%)**

Table 3 shows that RoLGaL consistently outperforms the baseline across all tested noise levels. The relative AUC lift increases from +0.31% at $\eta = 0.01$ to +0.43% at $\eta = 0.10$, indicating that the framework remains effective as injected label noise becomes more severe.

Parameter Tuning. To reduce tuning bias, we use the same learning rate of $1e-3$ across all models and fix RoLGaL hyperparameters across experiments. Unless otherwise specified, we set $\lambda = 1$ and use fixed diffusion variance and guardrail threshold configurations for each head.

5.4 Component Ablation Study

We conduct ablation studies to understand the contribution of each RoLGaL component under injected label noise. Table 4 reports results at $\eta = 0.01$, covering multi-head scaling and the contribution of the guardrail and label diffusion modules.

Increasing the number of heads from $K = 1$ to $K = 2$ improves the AUC lift from +0.27% to +0.31% and reduces variance. Using $K = 3$ achieves similar performance, suggesting diminishing returns from additional heads in this setting. Removing the guardrail causes the largest performance drop, reducing the lift to +0.12%, confirming that reference-guided correction is the most important component. Removing label diffusion also reduces performance, showing that soft-label correction complements guardrail-based filtering.

5.5 Online Production Validation

To evaluate practical impact, we deployed RoLGaL in a large-scale ads delivery system and compared it against the production baseline in an online A/B test.

Table 4: Component Ablation ($\eta = 0.01$). We report AUC Mean ± Std and relative lift.

Configuration	AUC (Mean ± Std)	Lift (Δ%)
A. Multi-Head Scaling (Full RoLGaL)		
Baseline	0.8989 ± 0.0002	0.00%
RoLGaL ($K = 1$)	0.9013 ± 0.0002	+0.27%
RoLGaL ($K = 2$)	**0.9016 ± 0.00005**	**+0.31%**
RoLGaL ($K = 3$)	0.9016 ± 0.00008	+0.30%
B. Module Contribution (Fixed $K = 2$)		
No Guardrail	0.8999 ± 0.0001	+0.12%
No Label Diffusion	0.9015 ± 0.0001	+0.29%
Full RoLGaL	**0.9016 ± 0.00005**	**+0.31%**

Deployment Efficiency. RoLGaL is implemented as an auxiliary training task. During serving, the auxiliary head is discarded and only the optimized main head is retained. Therefore, RoLGaL introduces zero inference latency overhead.

Offline Production Validation. Before launch, RoLGaL was evaluated on a hold-out production dataset and achieved a **-0.29%** improvement in Normalized Entropy (NE) over the baseline. This provides a calibration-sensitive offline signal before online testing, while preserving the serving graph.

Online A/B Test. We conducted a 12-day online A/B test on global traffic. As shown in Table 5, RoLGaL achieved a statistically significant **+0.23%** lift on a primary revenue-related business metric ($p < 0.05$), while adding no inference latency.

Table 5: A/B Test Results (12-day window). RoLGaL achieves statistically significant gains over the baseline on a primary revenue-related business metric.

Metric Segment	Relative Lift
Global Traffic	+0.23% ± 0.18%

Broader Deployment. The online A/B test was conducted on a flagship production ranker and served as the initial validation of RoLGaL under live traffic. After this positive result, the framework was adopted across 10+ production CVR models spanning different conversion objectives, including purchases, app installs, and other lower-funnel events. This broader rollout suggests that the auxiliary-task design is not tied to a single model or objective and can be integrated into mature ranking systems while preserving serving-time efficiency.

Overall, the public experiments and production results provide complementary evidence: RoLGaL improves robustness under controlled noisy-label stress tests, remains practical under strict serving constraints, and generalizes across multiple production conversion objectives.

6 Conclusion and Future Work

We introduced RoLGaL, a robust continual training framework for CVR prediction under heterogeneous label noise. By combining label diffusion, guardrail-guided correction, and auxiliary-task training, RoLGaL improves representation robustness while preserving the primary serving target and adding zero inference-time overhead.

Public benchmark experiments show that RoLGaL improves offline ranking quality and is more robust than Label Smoothing and Knowledge Distillation under injected label noise. Production validation further demonstrates practical impact: RoLGaL achieves a statistically significant +0.23% lift in a primary revenue-related business metric in online A/B testing and has been adopted across 10+ production CVR models. Together, these results show that selective label correction can improve continual ads ranking systems under both controlled noisy-label settings and real production constraints.

Future Directions. RoLGaL can be extended to broader ads and recommendation settings. First, RoLGaL can be extended to CTR prediction models, where click labels are more abundant but still noisy due to accidental clicks, exploratory browsing, and position or presentation bias. Our additional Avazu experiment in Appendix A provides preliminary evidence that RoLGaL remains robust in a public CTR prediction setting. Second, future work can adapt label diffusion and guardrail filtering to pairwise or listwise objectives [11, 13], where one corrupted label may affect multiple comparisons. Third, for sequential recommendation models such as SASRec [8], RoLGaL could provide reliability-aware attention masks that downweight noisy historical interactions. Finally, future work will evaluate RoLGaL under more structured attribution-delay noise, delayed-feedback baselines, and richer offline calibration diagnostics.

References

[1] Avazu. 2014. Click-Through Rate Prediction. https://www.kaggle.com/c/avazu-ctr-prediction. Kaggle competition dataset.

[2] Haoyan Chua, Yingpeng Du, Zhu Sun, Ziyan Wang, Jie Zhang, and Yew-Soon Ong. 2024. Unified Denoising Training for Recommendation. In *Proceedings of the 18th ACM Conference on Recommender Systems (RecSys)*. ACM.

[3] Diemert Eustache, Meynet Julien, Pierre Galland, and Damien Lefortier. 2017. Attribution Modeling Increases Efficiency of Bidding in Display Advertising. In *Proceedings of the AdKDD and TargetAd Workshop, KDD, Halifax, NS, Canada, August, 14, 2017*. ACM.

[4] Xiaoqiang Gui, Yueyao Cheng, Xiang-Rong Sheng, Yunfeng Zhao, Guoxian Yu, Shuguang Han, Yuning Jiang, Jian Xu, and Bo Zheng. 2023. Calibration-compatible Listwise Distillation of Privileged Features for CTR Prediction. In *arXiv preprint arXiv:2312.08727*.

[5] Bo Han, Quanming Yao, Xingrui Yu, Gang Niu, Miao Xu, Weihua Hu, Ivor Tsang, and Masashi Sugiyama. 2018. Co-teaching: Robust Training of Deep Neural Networks with Extremely Noisy Labels. In *NeurIPS*. 8527–8537.

[6] Zhuangzhuang He, Yifan Wang, Yonghui Yang, Peijie Sun, Le Wu, Haoyue Bai, Jinqi Gong, Richang Hong, and Min Zhang. 2024. Double Correction Framework for Denoising Recommendation. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. ACM.

[7] Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. 2015. Distilling the Knowledge in a Neural Network. *arXiv preprint arXiv:1503.02531* (2015).

[8] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *Proceedings of the IEEE International Conference on Data Mining (ICDM)*. 197–206.

[9] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. 2017. Overcoming catastrophic forgetting in neural networks. In *PNAS*.

[10] Junnan Li, Richard Socher, and Steven C.H. Hoi. 2020. DivideMix: Learning with Noisy Labels as Semi-supervised Learning. In *International Conference on Learning Representations*.

[11] Zhutian Lin, Junwei Pan, Shangyu Zhang, Ximei Wang, Xi Xiao, Shudong Huang, Lei Xiao, and Jie Jiang. 2024. Understanding the Ranking Loss for Recommendation with Sparse User Feedback. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24)*. ACM, 5409–5418.

[12] David Lopez-Paz and Marc'Aurelio Ranzato. 2017. Gradient episodic memory for continual learning. In *NIPS*.

[13] Boxiang Lyu, Zhe Feng, Zachary Robertson, and Sanmi Koyejo. 2023. Pairwise Ranking Losses of Click-Through Rates Prediction for Welfare Maximization in Ad Auctions. In *Proceedings of the 40th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 202)*. PMLR, 23239–23263. https://proceedings.mlr.press/v202/lyu23b.html

[14] Giorgio Patrini, Alessandro Rozza, Aditya Krishna Menon, Richard Nock, and Lizhen Qu. 2017. Making deep neural networks robust to label noise: A loss correction approach. In *CVPR*.

[15] Mengye Ren, Wenyuan Zeng, Bin Yang, and Raquel Urtasun. 2018. Learning to reweight examples for robust deep learning. In *International Conference on Machine Learning (ICML)*. 4334–4343.

[16] Jun Shu, Qi Xie, Lixuan Yi, Qian Zhao, Sanping Zhou, Zongben Xu, and Deyu Meng. 2019. Meta-Weight-Net: Learning an Explicit Mapping For Sample Weighting. In *NeurIPS*.

[17] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. 2016. Rethinking the Inception Architecture for Computer Vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2818–2826.

[18] Ruoxi Wang, Bin Fu, Gang Fu, and Mingliang Wang. 2021. DCN V2: Improved deep & cross network and practical lessons for web-scale learning to rank systems. In *Proceedings of the Web Conference 2021*. 1785–1797.

[19] Jieming Zhu, Jinyang Liu, Weiqi Li, Jincai Lai, Xiuqiang He, Liang Chen, and Zibin Zheng. 2020. Ensembled CTR Prediction via Knowledge Distillation. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM)*. ACM.

A Additional Results on Avazu

To further evaluate whether RoLGaL generalizes beyond the Criteo Attribution dataset, we conduct an additional experiment on the Avazu CTR Prediction dataset [1]. Although Avazu is a CTR dataset rather than a CVR benchmark, it provides a different public ads prediction setting with distinct feature and noise characteristics. We use a streaming-style setup with injected label noise ($\eta = 0.10$) and compare RoLGaL against Knowledge Distillation (KD), the strongest baseline from the clean Criteo experiment.

Table 6: Additional Results on Avazu with Injected Noise ($\eta = 0.10$). RoLGaL maintains robustness on a different public ads dataset.

Configuration	AUC (↑)	Lift (Δ%)
Baseline	0.7536 ± 0.0004	–
Knowledge Distillation	0.7535 ± 0.0010	-0.01%
Multihead RoLGaL ($K = 2$)	**0.7543 ± 0.0007**	**+0.10%**

As shown in Table 6, RoLGaL maintains a positive AUC lift on Avazu under injected noise, while KD slightly underperforms the baseline. This result is directionally consistent with the Criteo noise experiments: teacher-based stabilization can be competitive when labels are relatively clean, but it is less effective when explicit label corruption is introduced. Although Avazu is not a CVR benchmark, the result provides supporting public evidence that RoLGaL-style robust training can extend to another ads prediction task. We therefore view Avazu as complementary validation, while leaving broader cross-task evaluation for future work.