

ReLoop: A Self-Correction Continual Learning Loop for Recommender Systems

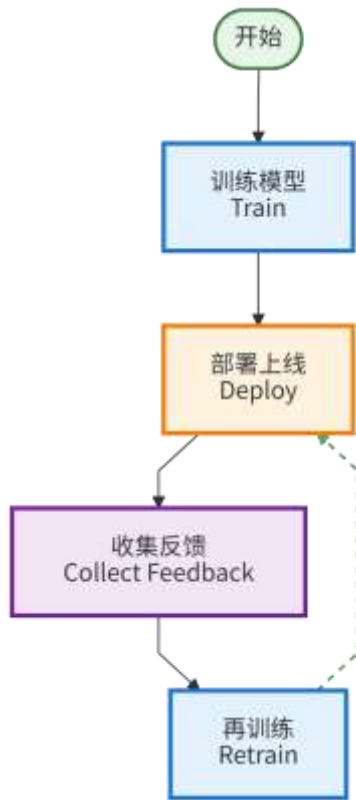
Guohao Cai, Jieming Zhu, et al. (Huawei Noah's Ark Lab)



徐凌飞 2025.11.10

研究背景

- 推荐系统广泛应用于电商、短视频、资讯等场景
- 模型需要频繁更新以捕捉用户兴趣变化
- 常规流程：训练 → 部署 → 收集反馈 → 再训练



持续学习在推荐系统中的挑战

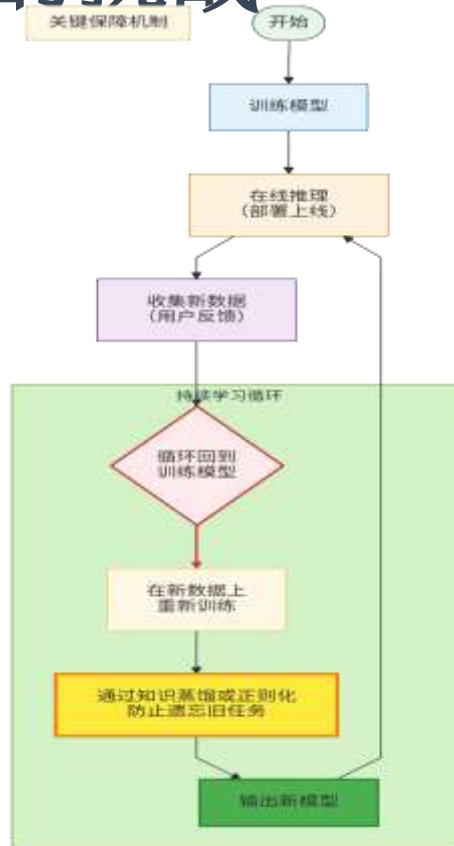
- 灾难性遗忘**：模型在学习新数据时丢失旧知识

- 缺乏自我纠错**：无法从过去错误中改进

ReLoop的思路：

- 从“记忆旧知识”转向“反思旧错误”

- 让推荐系统具备自我纠错的持续学习闭环



持续学习在推荐系统中的挑战

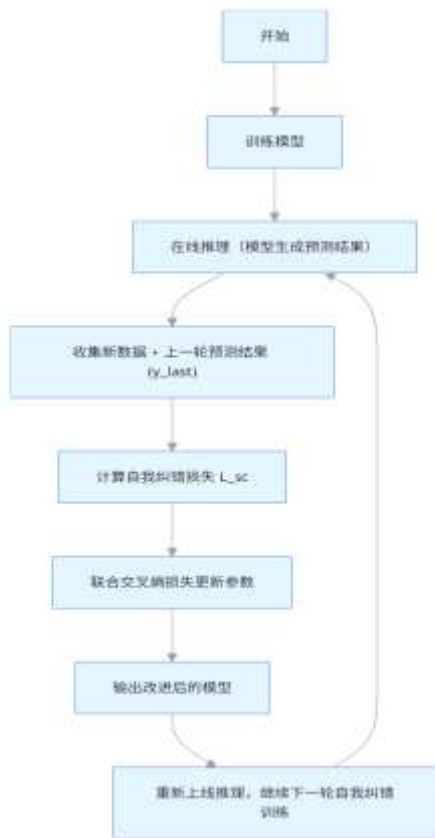
- **灾难性遗忘**：模型在学习新数据时丢失旧知识

- **缺乏自我纠错**：无法从过去错误中改进

ReLoop的思路：

- 从“记忆旧知识”转向“反思旧错误”

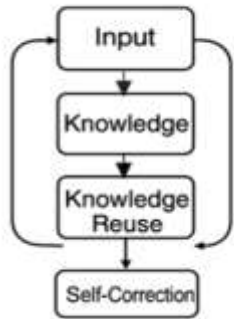
- 让推荐系统具备自我纠错的持续学习闭环



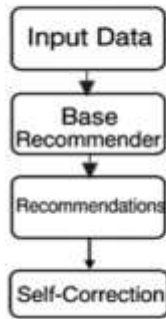
问题：缺乏自我纠错机制

- 模型版本之间相互独立
- 没有利用过去预测错误的机会
- 类比人类学习：人会复盘错题，机器不会

Human Learning

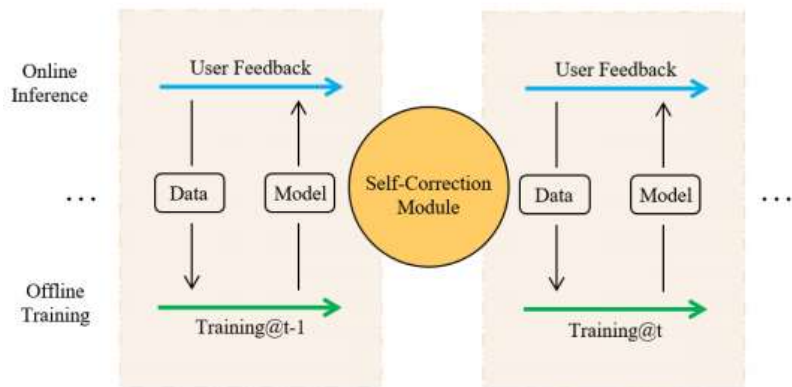


Recommender System



推荐系统的常规训练闭环

- 收集用户点击日志
- 模型训练CTR预测
- 部署上线推理
- 收集新数据再训练



(a) Self-Correction Loop

每轮模型训练都从头开始，不知道上一轮哪里做得不好。

研究动机与贡献

- 核心思想**: 让推荐系统学会“从错误中学习”

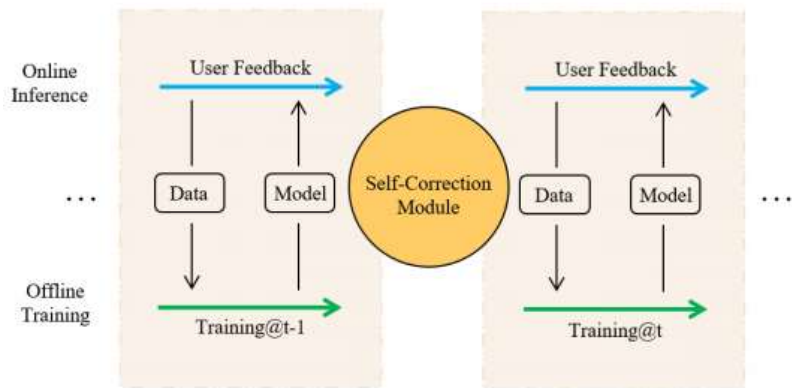
- 提出方法**: ReLoop 自我纠错持续学习框架

- 主要贡献**:

1. 通用的自我纠错机制

2. 新损失函数 Self-Correction Loss

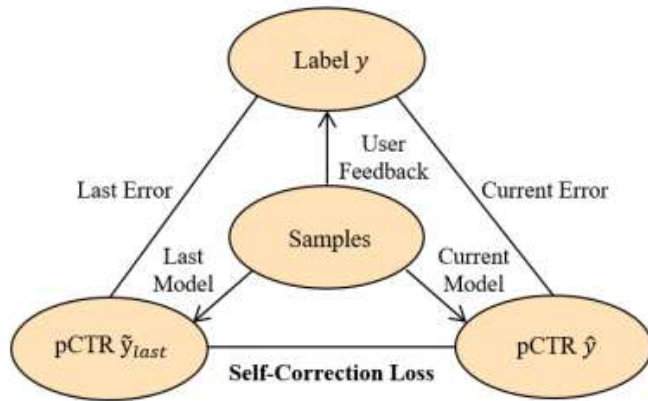
3. 离线与在线实验验证有效性



(a) Self-Correction Loop

ReLoop 框架整体架构

- 在两次训练之间加入自我纠错模块
- 新模型在训练时参考旧模型的错误

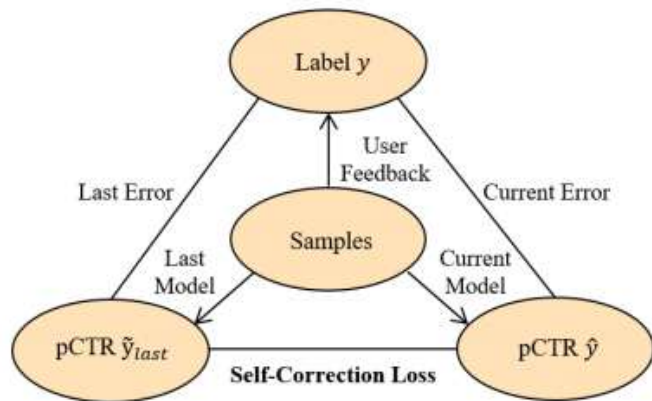


(b) Self-Correction Module

自我纠错模块原理

- 上一轮预测结果: \tilde{y}_{last}
- 当前预测: \hat{y} , 真实标签: y
- 若当前误差 > 上一轮误差 → 加惩罚

$$\text{即 } y - \hat{y} > y - \tilde{y}_{last} \Leftrightarrow \tilde{y}_{last} - \hat{y} > 0$$



(b) Self-Correction Module

自我纠错损失函数

自我纠错损失 L_{sc}

作者设计了一个类似于 **hinge loss** 的扩展版本：

$$L_{sc} = y \max(\tilde{y}_{last} - \hat{y}, 0) + (1 - y) \max(\hat{y} - \tilde{y}_{last}, 0)$$

y ：样本标签（正样本 = **1**，负样本 = **0**）

\hat{y} ：当前模型预测的点击率（**CTR**）

\tilde{y}_{last} ：上一版本模型预测的点击率

自我纠错损失函数

传统交叉熵损失 L_{ce}

常见的点击率预测任务目标函数：

$$L_{ce} = -y \log \hat{y} - (1 - y) \log(1 - \hat{y})$$

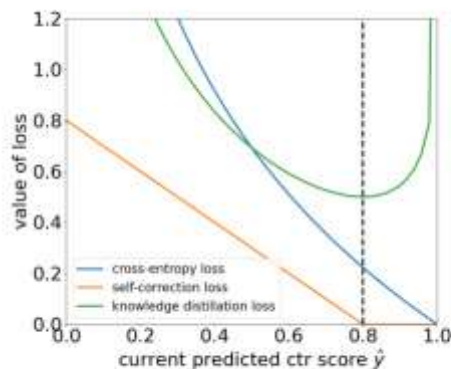
它保证模型具备基本的监督学习能力。

最终 ReLoop 损失函数结合两部分：

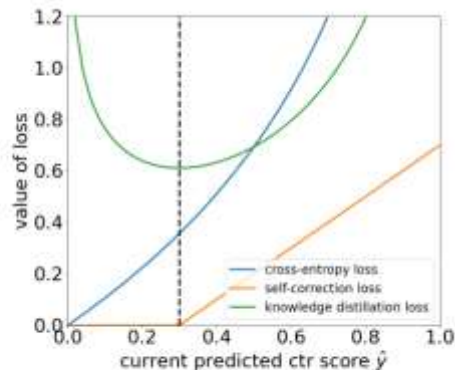
$$L = \alpha L_{sc} + (1 - \alpha) L_{ce}$$

与知识蒸馏的区别

- KD: 模仿教师 → 可能复制错误
- ReLoop: 纠正教师错误 → 主动超越



(a) $y = 1, \bar{y}_{last} = 0.8$



(b) $y = 0, \bar{y}_{last} = 0.3$

数据集与实验设置

- 公共数据集：Criteo、Avazu、MovieLens、Frappe
- 工业数据集：新闻推荐（500M样本）
- 指标：AUC、Logloss
- 对比模型：LR, FM, DeepFM, DCN, AFN+, KD

数据集类型	数据集名称	样本数量（约）	特征数量	特征类型	应用场景
公共数据集	Criteo	45M	39	数值 + 类别特征	在线广告点击率预测（CTR）
	Avazu	40M	22	类别特征为主	移动广告点击率预测
	MovieLens	2M	7	用户 + 电影特征	电影推荐
	Frappe	288K	10	上下文特征（时间、地点等）	移动应用推荐
工业数据集	News Recommendation	500M	100+	用户画像 + 内容特征	实际新闻推荐系统（在线实验）

数据集与实验设置

- 公共数据集：Criteo、Avazu、MovieLens、Frappe
- 工业数据集：新闻推荐（500M样本）
- 指标：AUC、Logloss
- 对比模型：LR, FM, DeepFM, DCN, AFN+, KD

类别

模型

说明

传统模型

LR, FM

基线模型（线性与二阶交互）

深度模型

DeepFM, DCN, AFN+

主流CTR模型

知识蒸馏模型

KD

引入教师模型指导学习

本文模型

ReLoop

自我纠错持续学习框架（可插拔模块）

离线实验结果

ReLoop 在 AUC、Logloss 等指标上均显著优于传统方法，表现出更好的稳定性和泛化能力。

Models	Criteo		Avazu		MovieLens		Frappe	
	AUC	Logloss	AUC	Logloss	AUC	Logloss	AUC	Logloss
LR	0.7858	0.4636	0.7313	0.4065	0.9215	0.3080	0.9329	0.2860
FM	0.7933	0.4574	0.7496	0.3740	0.9388	0.2797	0.9641	0.2143
NFM	0.7968	0.4537	0.7531	0.3761	0.9441	0.3004	0.9727	0.2079
IPNN	0.8026	0.4509	0.7526	0.3737	0.9469	0.2792	0.9735	0.2012
Wide&Deep	0.8062	0.4453	0.7529	0.3744	0.9381	0.3310	0.9728	0.2038
DCN	0.8059	0.4463	0.7550	0.3721	0.9419	0.2791	0.9402	0.2808
DeepFM	0.8025	0.4501	0.7535	0.3742	0.9424	0.3131	0.9719	0.2108
xDeepFM	0.8070	0.4443	0.7535	0.3737	0.9448	0.2717	0.9738	0.2098
FmFM	0.8056	0.4462	0.7603	<u>0.3685</u>	0.9465	0.2714	0.9749	0.2004
AFN+	0.8074	0.4451	0.7555	0.3718	0.9500	0.2585	0.9783	0.1762
KD+DCN	<u>0.8136</u>	<u>0.4473</u>	<u>0.7653</u>	0.3906	0.9648	<u>0.2164</u>	<u>0.9837</u>	0.1672
KD+DeepFM	0.8135	0.4462	0.7652	0.3690	<u>0.9650</u>	0.2183	0.9803	<u>0.1552</u>
RLP+DCN	0.8138	0.4381	0.7648	0.3667	0.9693	0.2189	0.9843	0.1660
RLP+DeepFM	0.8139	0.4382	0.7659	0.3750	0.9687	0.2160	0.9848	0.1549

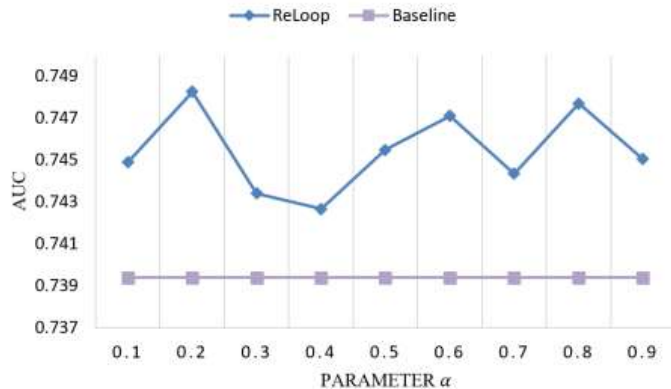
工业数据集与参数敏感性

- 部署在一个工业新闻推荐数据：AUC+1.2%， Logloss -21.8%

- α 参数敏感性 (Figure 3) ：性能稳定

- 训练开销+1.8%，推理无增加

	AUC	Improv.	Logloss	Improv.	Time cost
Baseline	0.7394	-	0.2959	-	-
KD+Baseline	0.7458	+0.87%	0.2743	+7.30%	+1.8%
RLP+Baseline	0.7483	+1.20%	0.2313	+21.83%	+1.8%



总结与反思

本研究提出了 **ReLoop 框架** —— 一种具备自我纠错能力的持续学习机制，用于推荐系统的模型训练。

不同于传统的持续学习仅关注“防止遗忘”，ReLoop 通过引入自我纠错损失函数，让模型能够**主动反思并修正过往预测错误**，形成一个**持续改进的学习闭环**。

展望与启示

未来的研究可以从更广泛和更精细的角度拓展 ReLoop 框架。

ReLoop 的自我纠错机制可以与 **记忆模块**、**自适应权重机制** 等方法结合，进一步增强模型自我进化能力。

这一框架不仅适用于推荐系统，还可以推广到其他需要持续优化的任务场景，如排序学习、在线广告投放或强化学习。总体而言，ReLoop 的思想为持续学习提供了一种新的视角。让模型在不断更新中学会“反思与改进”，迈向真正具备自我优化能力的智能系统。

The slide features a light blue background with decorative hexagonal shapes in the corners. The top-left corner has a cluster of overlapping hexagons in shades of blue and cyan. The top-right corner has a similar cluster. The bottom-left and bottom-right corners have single, partially visible hexagons in blue and cyan respectively.

请各位老师批评指正！