

申请上海交通大学博士学位论文

实时竞价广告中的建模与决策优化

论文作者 任 侃

学 号 0140339030

导 师 俞勇教授

专 业 计算机科学与技术

答辩日期 2019 年 5 月 20 日

Submitted in total fulfillment of the requirements for the degree of Doctor  
in Computer Science

# Modeling and Decision Optimization in Real-time Bidding Advertising

KAN REN

Advisor

Prof. YONG YU

DEPART OF COMPUTER SCIENCE AND ENGINEERING,  
SCHOOL OF ELECTRONIC INFORMATION AND ELECTRICAL ENGINEERING  
SHANGHAI JIAO TONG UNIVERSITY  
SHANGHAI, P.R.CHINA

May. 20th, 2019

# 上海交通大学 学位论文原创性声明

本人郑重声明：所提交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：\_\_\_\_\_

日期：\_\_\_\_\_年\_\_\_\_月\_\_\_\_日

# 上海交通大学 学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权上海交通大学可以将本学位论文的全部或部分内 容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于

保 密 ，在 \_\_\_\_\_ 年解密后适用本授权书。

不保密 。

(请在以上方框内打 )

学位论文作者签名： \_\_\_\_\_

指导教师签名： \_\_\_\_\_

日 期： \_\_\_\_\_年 \_\_\_\_月 \_\_\_\_日

日 期： \_\_\_\_\_年 \_\_\_\_月 \_\_\_\_日

# 实时竞价广告中的建模与决策优化

## 摘要

随着信息技术的发展，在线广告，即数字营销，已经成为电商网站、搜索引擎、社交媒体等互联网商业平台的主要收入来源。在线广告的目的则是通过网络媒体，构建商家与普通消费者之间的信息桥梁，为商家推广产品从而获取客源，为互联网用户解决自身商品需求提供合适的商品信息。为了达到在线广告的目标，基于数字化手段的计算广告研究依托于信息时代日益增长的算力资源，为线上数字营销提供了绝佳的解决方案。计算广告相关技术更是具备高扩展性、可计算及可优化等特性。

自从 2009 年首次提出以来，实时竞价广告已然成为在线广告领域中最被看好也是最普遍应用的数字营销范式。实时竞价广告使得广告主能够在实时的情形下，单独采购任一互联网用户的浏览行为所产生的流量与广告展示机会。一方面，实时竞价广告能够汇聚丰富的媒体与广告位资源，更能支持广告主采用更为灵活的线上营销策略，例如以广告投放的购买转化效果为目的，或以提高品牌认知度为投放目标。另一方面，它还能通过可计算的方式赋予广告主动态的用户定向机制，即针对不同流量、不同用户实施多样的广告展示策略。

在这个前沿研究领域，众多研究者已经针对上述几个基础研究课题及其他不断涌现的研究挑战提出了众多的解决方案。从技术的视角来说，这些解决方案涵盖了一系列广泛的研究领域，包括信息检索、机器学习、最优化理论、经济学与博弈论等。

在本文中，我们从一个综合的建模视角来剖析实时竞价广告中的几个主要的研究难题，并同时希望能够解决广告生态系统中的其他重要挑战。我们针对实时竞价广告中的几个任务进行了形式化定义，并分析了已有方法的缺陷，最终从以下五个角度提出了我们的解决方案。

**成本估计** 一般由竞价形势预估任务完成。对于竞价形势预估，即市场价格分布的建模，我们采用了一种基于生存分析方法的建模思想，并创新性地提出了一种自回归深度学习模型，建模竞价过程中的市场价格概率密度，同时解决数据偏差等问题。这也是业界首个考虑数据特征中时序信息的模型。该模型无需针对竞价形势的概率分布做一些理想化的数学假设，并能以灵活的形式建模、预测复杂的真实市场价格分布。

**效益估计** 则由用户反馈预估去解决。现有的解决方案一般将用户反馈预估看作是二分类机器学习任务，但在实时竞价广告场景中这并非最优建模方案。我们采取了一个新的优

化视角并提出了一种不同于传统模型的优化目标函数，直接通过反馈预估模型优化广告主的收益。经过推导得到的模型能够在优化反馈预估的分类效果的同时，考虑投放成本与竞拍效果等，因而在实际场景中显著提高了广告投放效益，节省了广告预算。

**竞拍优化** 常常被研究者作为一个独立的模块，与用户反馈预估等作为单独的任务进行分别优化。然而，效益预估、成本预估及竞拍优化之间存在相辅相成的紧密联系，因此在解决其他问题的同时考虑竞价优化事实上是实时竞价广告等场景中一种更为有效的综合优化方式。我们采用了一种新型的目标函数，以广告投放收益为优化目标，通过一种统一的优化框架，联合优化效益预估、成本预估与竞价策略。该框架在实际场景中显著提高了广告投放的效果。

**用户建模** 无论在实时竞价广告还是其他线上智能信息系统中，都是一个重要的基础模块。它不仅能够刻画用户行为背后的兴趣与行为模式，更能为后续用户行为预测提供重要的线索，也就能给信息系统的决策提供重要的思路，例如提升广告投放的效果，推进推荐系统的性能等。我们针对线上平台丰富而又多样的用户行为提供一种终生用户建模框架，通过层级的周期记忆网络建模、记忆与存储用户在不同时期、不同场景中的行为模式，为任何时候用户行为预测提供了一种综合解决方案。

**用户转化归因分析** 是在线广告系统运行过程中一个重要的提效手段，它能够指导后续广告投放方案，例如更合理的预算分配计划。为了达成这一目标，我们基于深度神经网络构建了一种多触达转化归因分析方案。这个神经网络模型不仅大幅提高了用户转化的预测准确率，更能提供一种更加合理的转化归因分析思路，用以探索广告投放过程中，广告内容的触达对用户最终购买等转化行为的影响。为了量化地评估不同归因方法的效果，我们还提供了一种兼具新颖性与实用性的离线评估框架，通过评估不同归因模型指导下预算分配方案的合理性，继而对比不同归因方法的有效性。

另外，本文所述方法均在真实场景数据集或在线广告平台上验证了方法的有效性，获得了学术界与工业界的认可。以上这些研究课题不仅能在实时竞价广告场景中解决实际的问题，更能在多个相关领域得到广泛的应用，例如推荐系统与搜索引擎中的排序与筛选优化、经济学中的竞价机制优化，等等。

**关键词：** 机器学习 实时竞价 计算广告 用户反馈预估 用户建模 决策优化 竞价形势预估

# Modeling and Decision Optimization in Real-time Bidding Advertising

## ABSTRACT

Nowadays, online advertising, a.k.a. digital marketing, whose goal is to bridge the gap between merchandise sellers and common customers through the online media, has become the main source for the Internet-based business platform, such as e-commerce websites, search engines and social media. To achieve the goal of digital marketing, computational advertising provides a programmatic approach with scalable, computational and even learnable methodologies via rapidly increasing computing capability.

Emerged from 2009, real-time bidding (RTB) is the most widely adopted online advertising paradigm which essentially facilitates buying an individual ad impression in *real time* while it is still being generated from a user's visit. On one side, RTB scales up the online buying process by aggregating abundant media inventories and supporting flexible marketing strategies. On the other side, it also enables dynamic user targeting via computational approaches.

In this cutting-edge frontier area, many researchers have devoted lots of efforts to solve the emerging challenges from these three fundamental aspects of RTB, along with the other considerations of online advertising. From the view of techniques, the solutions cover a wide range of research fields such as information retrieval, machine learning, optimization, economics and game theory.

In this thesis, we take a comprehensive modeling view to tackle the main problems of RTB and deal with several other challenges in the RTB ecosystem. We define the tasks of modeling in RTB advertising, and analyze the problems hidden in the existing methods, and finally propose our solutions from five perspectives as bellow.

**Cost Estimation** For bid landscape forecasting, we proposed an auto-regressive deep model which is the first work considering sequential feature dependency with survival analysis for modeling market price distributions. Thus it does not need to make any assumptions about base distribution forms of bid landscape and can generally predict flexible market price distributions.

**Utility Prediction** The existing methods for user response prediction usually regard the problem as a binary classification task, which may not be optimal in RTB scenario. We propose a new objective function for directly optimizing advertiser profits when conducting ad campaigns. The derived model has significantly improved the advertising utility since our method is cost-sensitive which can largely save the budget and allocate for high-effective ad impressions with low costs.

**Bid Optimization** The researchers often optimize the bid function and user response prediction function w.r.t. the isolated objectives. They have never combined these two problems. However, it is more effective considering the overall optimization procedure in RTB scenario. Thus it is eager to optimize through a comprehensive approach to improve the model efficiency. We propose a unified objective function of advertiser's profits and optimize it from three aspects, i.e., user response prediction, bid landscape forecasting and bid optimization, in the whole framework, which has significantly improved the advertising performance.

**User Modeling** In RTB or other online information systems, user modeling is a key component for better user interest capturing and behavior pattern mining. Thus it provides some valuable clues for subsequent decision making in information systems, such as better advertising and recommendation, etc. To model the abundant yet multi-facet user behaviors in the online system, we propose a lifelong user modeling framework with hierarchical and periodic memory network, to store the user behavior patterns along her long-term history of online activities. It also conducts a comprehensive solution for user behavior prediction.

**Conversion Attribution** To better understand user behaviors and direct the subsequent ad delivery, we also conduct a deep neural network for multi-touch conversion attribution. The proposed deep model improves the conversion prediction in a large margin and it also provides comprehensive conversion attributions for each user behavior sequence. To quantitatively benchmark attribution models, we also propose a novel yet practical attribution evaluation scheme through the proxy of budget allocation (under the estimated attributions) over ad channels.

Besides, both the academic community and industrial applications have demonstrated the effectiveness and efficiency of the modeling perspective and the corresponding solutions in this thesis. We not only evaluate our models against state-of-the-arts on several large-scale datasets, but also conduct online experiments on real-world advertising platforms. Moreover, the pro-

posed methodologies can also facilitate many more widely spread applications in other research areas such as recommender systems and search engines.

**KEY WORDS:** Machine Learning, Real-time Bidding, Computational Advertising, User Response Prediction, User Modeling, Bid Landscape Forecasting, Bid Optimization, Functional Optimization, Conversion Attribution

# 目 录

插图索引	x i
表格索引	x ii
算法索引	x iii
主要符号对照表	x iv
<b>第一章 绪论</b>	<b>1</b>
1.1 研究背景	1
1.1.1 数字营销与计算广告	2
1.1.2 实时竞价广告	4
1.2 研究内容	5
1.3 研究意义	7
1.4 章节安排	8
<b>第二章 市场建模与成本预估</b>	<b>10</b>
2.1 二价交易机制	10
2.2 竞价形势预估	11
2.3 相关工作	13
2.3.1 竞价形势预估	13
2.3.2 删失问题处理	14
2.3.3 循环神经网络	15
2.4 基于深度生存分析的方法	15
2.4.1 问题定义	15
2.4.2 模型介绍	16
2.5 实验分析	21
2.5.1 数据集与实验流程	21
2.5.2 评估指标与对比模型	22
2.5.3 实验结果与对比分析	23

2.6	本章小结	26
<b>第三章</b>	<b>用户建模与效益预估</b>	<b>28</b>
3.1	用户反馈预估	28
3.2	相关工作	30
3.2.1	用户反馈预估	30
3.2.2	竞价策略	30
3.3	直接优化广告收益的方法	31
3.3.1	问题定义	31
3.3.2	模型介绍	32
3.3.3	模型分析	36
3.4	实验分析	38
3.4.1	数据集与实验流程	39
3.4.2	评估指标与对比模型	39
3.4.3	实验结果与对比分析	40
3.4.4	在线实验	44
3.5	本章小结	45
<b>第四章</b>	<b>竞价决策优化</b>	<b>47</b>
4.1	竞价策略优化	47
4.1.1	竞价机器模型	48
4.2	基于广告效益的联合优化	49
4.2.1	针对用户反馈预估的效益优化	49
4.2.2	针对竞价形势预估的效益优化	49
4.2.3	竞价策略优化	52
4.2.4	竞价机器优化算法	53
4.3	实验分析	54
4.3.1	实验流程与对比模型	54
4.3.2	实验结果与对比分析	55
4.3.3	在线实验	57
4.4	本章小结	58
<b>第五章</b>	<b>终生用户建模</b>	<b>60</b>
5.1	用户行为建模	60

5.2	相关工作	62
5.2.1	序列化用户建模	62
5.2.2	记忆神经网络	63
5.3	多层级周期记忆网络	63
5.3.1	符号与定义	63
5.3.2	终生序列建模	64
5.3.3	模型结构	65
5.3.4	预测与损失函数	69
5.4	实验分析	70
5.4.1	实验设置	70
5.4.2	实验结果与分析	74
5.4.3	可视化分析	76
5.5	本章小结	78
<b>第六章</b>	<b>转化归因</b>	<b>79</b>
6.1	多触达的转化归因	79
6.2	相关工作	82
6.3	深度转化归因方法	83
6.3.1	问题定义	83
6.3.2	序列建模	84
6.3.3	归因学习	86
6.3.4	离线归因评估框架	88
6.4	实验分析	90
6.4.1	数据集	90
6.4.2	实验流程与评估指标	91
6.4.3	对比模型	92
6.4.4	转化率预估实验性能	92
6.4.5	预算重分配离线评估	93
6.4.6	可视化分析	95
6.5	本章小结	98
<b>第七章</b>	<b>总结与展望</b>	<b>101</b>
<b>附录 A</b>	<b>竞价机器算法的在线更新公式</b>	<b>103</b>

---

附录 B 最优竞价策略理论证明	105
附录 C 最优竞价策略的博弈分析	107
C.1 基于收益优化的博弈分析 . . . . .	107
C.1.1 问题设定 . . . . .	107
C.1.2 对称博弈的最优竞价策略函数 . . . . .	108
C.1.3 “公地悲剧” . . . . .	111
附录 D 循环神经网络的实现	113
参考文献	114
致 谢	126
攻读学位期间发表的学术论文	127
攻读学位期间参与的项目	129

## 插图索引

1-1	品牌广告和促销广告	1
1-2	世界上首个在线横幅广告	2
1-3	实时竞价广告流程图	4
1-4	实时竞价广告的任务分解与本文研究内容安排	5
2-1	二价交易机制示意图	10
2-2	竞价形势预估任务	11
2-3	数据删失导致的幸存偏差图示	12
2-4	统计得到的真实市场价格分布与对应的胜率函数曲线	13
2-5	深度形势预估模型的框架图	17
2-6	胜率函数曲线的两种例子	19
2-7	模型学习曲线	25
2-8	结果可视化	26
3-1	期望效益模型的梯度分析	34
3-2	风险回报模型的梯度分析	35
3-3	用户反馈预估模型的学习曲线	42
3-4	模型输出的竞拍价格分布示意图	43
3-5	在线实验结果	45
4-1	竞价机器模型的联合优化框架图	48
4-2	竞价形式与胜率函数的几种闭型示意图	51
4-3	iPinYou 数据集上带预算约束的竞价性能展示	57
4-4	竞价机器算法不同模块的学习曲线	58
4-5	在线实验结果	59
5-1	某在线电商平台的用户点击行为统计 (2018 年 4 月至 9 月)	61
5-2	终生户序列建模的示意图	65
5-3	层级周期记忆网络 HPMN 的示意图	66
5-4	用户反馈预估的模型结构图	68
5-5	训练/测试数据集切分示意图	72

5-6	模型在不同数据集上的学习曲线 . . . . .	76
5-7	模型建模所得长短期及多跨度用户序列模式可视化 . . . . .	77
5-8	模型效果与层数及记忆槽数量的关系 . . . . .	78
6-1	用户浏览与广告触达示意图 . . . . .	80
6-2	用户浏览行为序列长度统计及平均转化率随不同序列长度的变化 . . . . .	82
6-3	采用双注意力机制的序列建模 . . . . .	85
6-4	注意力机制的计算过程 . . . . .	87
6-5	转化归因模型的学习曲线 . . . . .	94
6-6	转化归因指导的预算重分配评估实验结果 . . . . .	96
6-7	触达点角度的归因统计结果 . . . . .	97
6-8	Miaozhen 数据集上不同渠道的归因结果 . . . . .	98
6-9	Criteo 数据集上不同广告活动的归因结果 . . . . .	99
6-10	归因模型的归因结果计算所得的 ROI 分布 . . . . .	99
6-11	Criteo 数据集上的 $\lambda$ 结果 . . . . .	100

## 表格索引

2-1	竞价形势预估离线实验数据集统计	22
2-3	平均负对数似然概率 ANLP 评估结果	24
2-4	C 指数评估结果	24
3-1	两种用户反馈预估模型与传统方法的梯度对比	38
3-2	用户反馈预估模型的回归性能对比	41
3-3	广告活动利润效果对比 (基线模型)	43
3-4	广告活动利润提升效果对比 (相比基线 CE 模型)	46
3-5	其他离线实验指标效果对比	46
4-1	双模块优化与基线模型对比	56
4-2	不同预算下的广告投放收益详情	58
5-1	符号标记与含义描述	64
5-2	数据集信息统计	72
5-3	不同数据集上的 HPMN 模型结构超参数详细设定	74
5-4	性能对比结果	75
6-1	转化率预估实验的评估结果	93
6-2	预算重分配的离线评估实验结果	95

## 算法索引

4-1 竞价机器的迭代优化算法 . . . . .	53
6-2 预算重分配评估算法 . . . . .	90

## 主要符号对照表

$x$	竞价请求及广告内容特征向量
$y$	用户反馈标记
$z$	市场价格
$b$	广告主竞拍价格
$p(z)$	市场价格 (连续) 概率密度函数
$p_z$	市场价格 (离散) 分布列
$W(b)$	胜率函数
$S(b)$	输率函数
$b(\cdot)$	出价函数

## 第一章 绪论

本章将介绍数字营销 (Online Marketing) 及计算广告 (Computational Advertising) 方法的历史背景与算法机制变迁, 同时针对实时竞价广告生态系统进行多方面的介绍, 并从研究与实践两方面阐述实时竞价广告中的角色、系统与流程。

### 1.1 研究背景

广告 (Advertising), 或者称为市场营销 (Marketing), 是一种通过媒介传递或传播, 使得目标个体或人群获知所要传递的信息, 并能促成后续目标行为的方式。一般情况下, 人们都将广告指代为狭义的“商业广告”, 即商家为了促进品牌与商业服务或其产品在客户或潜在客户中的认知与消费而发起的信息传播行为。广义上的广告还包括公益广告、竞选广告等形式, 即沟通信息、促进认识的广告传播活动。这些广告形式均不作为商业推广行为, 不以营利为目的。

广告一般分为两种:

- 品牌广告 (Brand Ad): 提升商家品牌影响力, 促进大众的品牌认知, 无需直接提升产品或服务的销售, 如图1-1(a)所示;
- 促销广告 (Promotion Ad): 商家直接为了提升产品或服务的销售而做的广告, 能



(a) Brand Ad



(b) Promotion Ad

图 1-1 品牌广告和促销广告  
Fig 1-1 Brand ad and promotion ad

够直接造成用户的转化（例如点击、购买等行为）或商品的销售，如图1-1(b)所示。

广告需要通过媒介进行传播。自从广告诞生以来，纸质媒体、电视、广播和广告牌是主要的几种传统媒介。随着互联网的出现，伴随着基于互联网的商业模式的形成，人们获取信息的方式出现了极大的革新。互联网用户能够利用网络表达他们的信息诉求，例如搜索网页获取所需信息；或者浏览特定的网站并进行电子化的信息交换或电子交易，例如在电商平台进行购物。其中，信息系统（Information System）作为连接数字内容与互联网用户的重要媒介，承担着匹配内容与浏览者的功能<sup>[1]</sup>。互联网商业模式及其面向大众的信息系统往往以“免费”、“开放”而著称，但商业公司发展离不开经济收益，互联网商业模式的发展伴随着在线广告（Online Advertising）的兴起，并以后者作为其重要的收入来源，即以广告投放所获得的收益作为支持互联网服务平台发展的重要经济支撑。广告以电子化的形式，并以其数字化等特点，越来越受到产业界与学术界的关注。广告行业进入了数字营销与计算广告时代。

### 1.1.1 数字营销与计算广告

第一个在线广告<sup>1</sup>出现在 1994 年，当时全球仅有 3000 万互联网用户。在 HotWired 杂志 1994 年 10 月 27 日这一期上，其在线版本刊登了来自美国电信运营商 AT&T 的首个真正意义上的在线横幅广告，如图1-2所示。这个广告以问题的形式吸引读者（互联网用户）注意力并引导他们进行下一步的反馈行为，例如点击。



图 1-2 世界上首个在线横幅广告

Fig 1-2 First true banner ad on the Internet

随着越来越多的商务与服务以电子化的形式呈现，数字营销正在成为一种新的广告范式。首先出现的第一种在线广告形式是“赞助搜索”（Sponsored Search）。赞助搜索最初由 Idealab 的 Bill Gross 于 1998 年发明，用于支持网站 goto.com。该网站在 2001 年更名为 Overture，并在 2003 年被美国雅虎（Yahoo!）公司收购<sup>[2]</sup>。同时，谷歌（Google）公司在 2002 年 2 月启动了他们的 AdWords 业务，并使用“一般二价拍卖”（Generalized Second Price Auction, GSP）的形式在谷歌搜索引擎上进行基于搜索关键词的广告位售卖<sup>[3]</sup>。在用户端，每当用户发起一次搜索，搜索引擎将根据搜索关键词匹配相应的广告

<sup>1</sup>来源：<http://thefirstbannerad.com/>

展示给用户；在广告主端，广告主根据自身的广告属性对特定的关键词进行出价拍卖。二价拍卖形式将在第二章进行介绍。在 2002 年 5 月，谷歌公司为其赞助搜索广告添加了基于广告效果的竞价体系。雅虎公司也在 2007 年为搜索广告添加了同样基于效果的竞价体系，即广告主针对广告效果，例如用户的点击等行为进行广告投放出价，而非过去仅仅根据广告展示进行计费。值得注意的是，相关报道<sup>[4]</sup>称谷歌公司曾给予雅虎公司总计 2700 万美元股权作为专利使用赔偿，因其使用了雅虎公司所拥有的部分专利，这些专利主要是基于用户搜索结果匹配相关广告内容的算法。

网页搜索目前已经成为了人们日常生活的一部分，它帮助互联网用户极大地减少了获取所需信息的难度与时间成本。与此同时，赞助搜索广告帮助广告主提前买下特定关键词相关的广告展示机会，并在用户使用搜索引擎时进行广告展示，以满足广告主的商务诉求，例如出售更多服务与商品。另外，赞助广告之所以称之为“赞助”（Sponsored），是因为其帮助搜索引擎以对用户免费的形式为广大互联网用户提供高质量的搜索服务，也是谷歌等搜索引擎公司赖以生存的商业基石。

另外一种在线广告形式是出现于 1998 年的“展示广告”（Display Advertising），伴随着一种名为上下文广告的概念<sup>[5, 6]</sup>。Gilad Elbaz 和 Adam Weissman 创办的 Oingo 公司最早提出了一种基于单词含义以及 WordNet<sup>1</sup>词典的广告匹配与搜索算法，并在 2003 年 4 月被谷歌收购，被后者更名为 AdSense。此后，雅虎、微软（Microsoft）等公司相继推出了类似的基于展示列表的广告服务。这种上下文广告平台的出现主要是为了适应更加丰富的多媒体网络环境，例如视频、音频和基于地理位置的移动网络。这些平台能帮助发行方（Publishers）通过售卖网页上未安置内容的区域、无意义的视频片段或其他空置资源来放置广告，并以此从广告主端获利。一般这些广告服务台被称为“广告网络”（Advertising Networks）或“展示网络”（Display Networks）。这些平台并不由搜索引擎控制，但是却能对接数量庞大的个体广告主与个体发行方。

不过，人们一般也会将赞助搜索广告看作是一种上下文广告，因为前者其实是将用户搜索关键词看作上下文。但赞助搜索广告因为发展较早，市场容量庞大，早已在业界获得了较多的关注与较为深入的研究。相比之下，上下文广告及其后续衍生出来的新的在线广告范式（即本文重点介绍的实时竞价广告）作为新的广告生态体系，正越来越受到重视。

同时，在线广告的数字化、规模化、范围广、用户反馈实时等特征使得人们对于计算机自动广告投放、效果追踪与目标人群定向（投放）产生了巨大的诉求，于是催生了基于在线广告业务的计算广告研究，也是后续我们所要展开的研究内容的一个产业化概念。本论文将在后续介绍基于可计算方法与人工智能相关的理论，并讨论他们在一种全

---

<sup>1</sup>WordNet: <https://wordnet.princeton.edu/>

新在线广告范式中的应用。

### 1.1.2 实时竞价广告

在 2005 年左右，一种基于实时竞价（Real-time Bidding, RTB）方式的购买、售卖广告展示（Impressions）的平台开始出现，例如 ADSDAQ、AdECN、DoubleClick、adBrite 以及 Right Media 等，统称为广告交易所（Ad Exchanges）。与传统广告网络不同的是，这些广告交易所汇聚了多个广告网络的资源，并构建了一种实时竞价的方式来售卖广告展示机会。这些广告展示机会来自于用户访问网页时，网页上尚未加载出具体内容的广告位<sup>[7]</sup>。其“实时”体现在整个过程几乎不可被用户察觉的快速运转特性；而“竞价”意为广告主通过竞拍形式参与广告位售卖的竞争，后文将会继续针对这两点详细展开。

同时，实时竞价广告也催生了两种分工不同的市场角色<sup>[8]</sup>：

- 需求方平台（Demand Side Platform, DSP）：用于服务广告主并代其购买用户流量、参与实时竞价以运营广告活动的平台；
- 供给方平台（Supply Side Platform, SSP）：用于服务发行方并代其将用户流量通过广告交易所售卖给各需求方的平台。

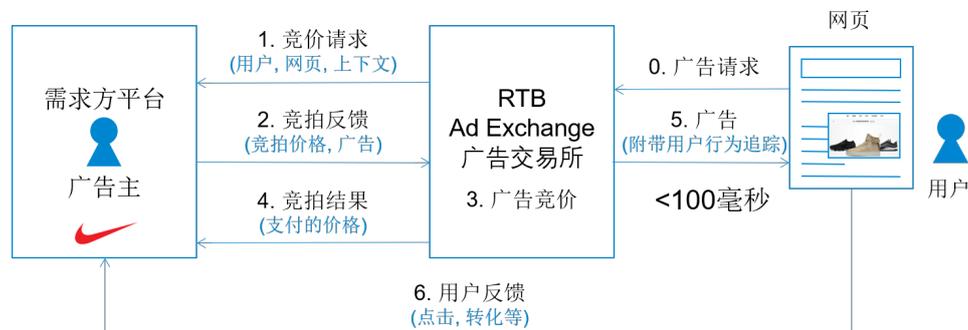


图 1-3 实时竞价广告流程图

Fig 1-3 Data Flow of Real-time Bidding Advertising

实时竞价广告的具体流程如图1-3所示，即

- (0) 用户在访问网页、浏览 APP 内容或者在社交媒体中进行互动时会产生广告请求。
- (1) 这个广告请求被供给方平台发送给广告交易所；广告交易所将其封装为竞价请求，即发起一个竞价交易，通知给所有对接的广告主代理平台并要求出价。
- (2) 广告主代理平台在经过内部评估、出价决策后给予出价反馈，即愿意以一定的成本买下这个广告展示机会。
- (3) 广告交易所根据二价竞价模型决定：出价最高的人赢得本次竞价，并支付“市场价格”（market price）或“赢得价格”（winning price)<sup>[9]</sup>，即赢家的所有竞争者中的最高价

格，亦即市场中的第二高价。

- (4) 赢得竞价的广告主收到竞拍获胜通知。
- (5) 该广告主将其广告内容呈现给用户端。
- (6) 用户反馈数据，包括用户点击、购买转化等行为将被记录并反馈给赢得竞价的广告主。

以上整个实施竞价流程（步骤 0 至步骤 5）每次在 100 毫秒内完成<sup>[7]</sup>，所以相对来说整个过程无法为用户所感知，谓之“实时”。

众多实时竞价广告系统正在不断产生大量的数据，每天有数以亿计的竞价交易在不同广告主的投放平台上进行。例如，广告主代理平台 Fiksu 每天处理着 480 亿的广告展示<sup>[10]</sup>；另一家知名广告主代理平台 Turn 也在曾汇报称达到了每秒最高 250 万广告展示的峰值流量<sup>[11]</sup>。为了更好地说明量化结果，我们对比了上海证券交易所每天的交易量约为 132 亿<sup>[12]</sup> 笔。可以说，基于展示广告的在线实时交易流量已经超过了金融市场。同时值得注意的是，基于实时竞价的展示广告产业为广大计算机科学家与经济学家提供了研究互联网流量、用户行为与动机以及在线交易相关课题的绝佳机会。这些交易与用户流量均为全球性质，更具有研究的一般性<sup>[13]</sup>。

## 1.2 研究内容

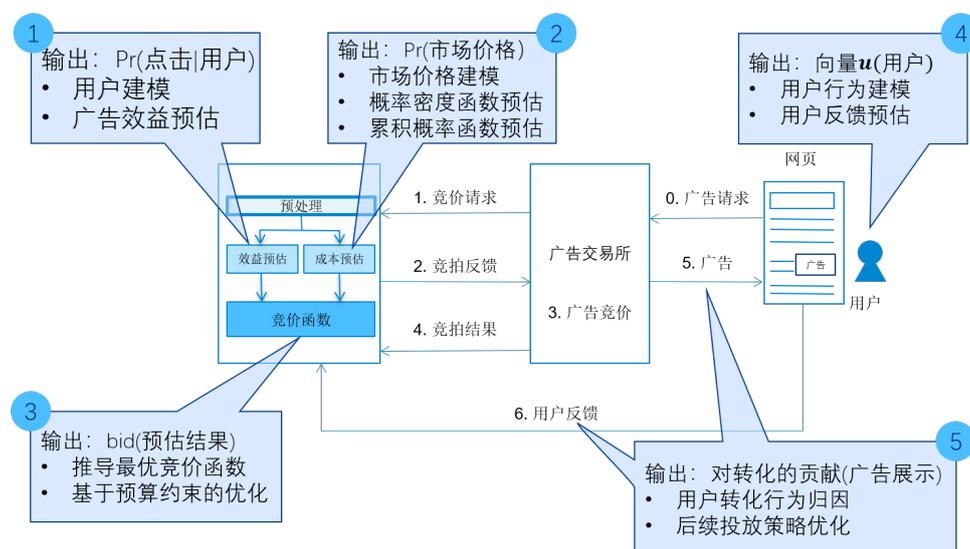


图 1-4 实时竞价广告的任务分解与本文研究内容安排

Fig 1-4 Tasks of Real-time Bidding and the main contents of this thesis

在线广告的根本目标是：在合适的时间以广告主和发行人认可的合适的价格，自动

推送合适的广告内容给合适的用户。

通过可计算 (Computational) 的形式, 并且在实时情况下, 实时竞价广告为用户和广告主提供了灵活而精准匹配的机会。从用户端考虑, 实时竞价广告能为用户提供个性化、高质量的广告资源, 能为每一位用户量身定做适合于其个体兴趣的广告展示, 真正做到“千人千面”; 从广告主端考虑, 实时竞价广告为广告主提供了实时触达全球互联网用户的可能性, 并且能为其提供灵活的广告投放策略, 包括基于基于用户建模的人群定向、用户广告个性化、广告投放与追踪自动化及广告预算智能分配等等。实时竞价广告还极大促进了基于人工智能的自动化算法来优化广告和用户之间的匹配 (Matching)。

图1-4展示了广告主需求方平台内部的预测决策策略流程。广告竞价请求 (Bid Request) 由上文提到的广告交易所转发给各个广告主代理平台。请求数据中包含了用户端信息 (例如浏览器信息, IP 信息)、广告页信息 (例如广告位规格) 以及上下文信息 (例如广告页面地址) 等<sup>[4]</sup>。在收到广告竞价请求后, 广告主会针对竞价请求数据进行一些预处理, 包括数据清洗 (包括过滤垃圾流量等)、特征工程、用户信息对齐 (将用户信息与 DSP 数据服务中的已有信息进行对齐) 等处理, 以供后续进行预测与决策任务。广告主会根据广告位与用户的信息进行效益估计, 来评估广告投放可能带来的收益, 例如用户关注与商品售卖。同时广告主还需要评估广告投放与展示的预期成本, 因为参与实时竞价会带来流量交易的经济开销。最后, 广告主会根据预估的效益与成本, 以及自身的预算情况进行竞价决策, 即决定最终参与实时竞价的竞拍价格。

从以上流程中我们可以看出, 广告主平台面临五方面主要的建模任务及挑战, 如图1-4所示:

**成本预估 (Cost Estimation)** 用于评估当前广告展示的市场价格分布。市场价格即赢得竞拍以后所支付的价格, 也就是广告展示的成本。

**效益预估 (Utility Estimation)** 用于预测用户反馈以预估广告投放的收益, 本质上是估计用户对于广告内容的正反馈 (点击、购买转化等) 概率。广告主只有在获得用户正反馈后, 才能达成广告的根本目的, 即用户关注与商品售卖。

**竞价优化 (Bid Optimization)** 建模出价策略并进行优化, 目标是在给定广告主预算下获取尽可能多的用户正反馈或转化 (conversions) 行为 (购买服务成为正式客户), 同时降低投放成本。

**用户建模 (User Modeling)** 在实时竞价广告还是其他线上智能信息系统中, 都是一个重要的基础模块。它不仅能够刻画用户行为背后的兴趣与行为模式, 更能为后续用户行为预测提供重要的线索, 也就能给信息系统的决策提供重要的思路, 例如提升广告投放的效果, 推进推荐系统的性能等。

**用户转化归因分析 (Conversion Attribution)** 用于分析广告投放过程中多次不同

的广告内容触达对于用户转化行为的影响。这是在线广告系统运行过程中一个重要的提效手段，它能够指导后续广告投放方案，例如更合理的预算分配计划。

本文将从以上五个研究课题入手，具体研究内容框图如图1-4所示。

从实时竞价广告研究所涉及的研究方法进行分类，解决方案基本涵盖了机器学习领域的几个典型问题：分类（Classification）、回归（Regression）、优化（Optimization）和决策（Decision Making），具体来说：

**分类** 针对用户正反馈的建模，通常定义为预测用户正反馈行为概率的二分类问题<sup>[15-17]</sup>，其中涉及到了逻辑斯蒂回归<sup>[18]</sup>、树模型<sup>[17]</sup>、因式分解机<sup>[19]</sup>等常用机器学习分类模型。

**回归** 预测用户的正反馈概率，有时候也会采用基于回归的方法来拟合用户的行为概率。同时，预测市场价格的分布往往需要使用基于回归方法的求解方案。本文的回归模型主要涉及到最大似然估计等统计模型。

**优化** 本文的优化方法涉及到两方面，一方面是基于实际问题提出新的目标函数，这部分将在第三章和第四章涉及到；另一方面则是针对已经建模好的问题，利用凸优化、泛函优化等数学方法进行求解，而这些数学方法将在全文的解决方案中贯穿始终。

**决策** 一般的人工智能系统除了预测之外，还需要在决策方面提供优化方案。我们通过优化决策函数的方法来解决实时竞价广告中的决策问题，即最优出价的决策。

同时，实时竞价广告还涉及到了经济学中的交易理论与博弈论，以及信息检索系统中数据库、排序与召回等技术，所以基于实时竞价广告的建模与优化是一种综合性的解决方案体系。

### 1.3 研究意义

基于计算机的大数据分析与智能决策系统不仅在科学研究等领域日益受到关注<sup>[13]</sup>，其在多个工商业及社会研究领域也拥有极为广泛的应用。其中，在线广告作为大数据分析的重要应用场景，呈现出研究与产业共同快速发展的繁荣局面。基于实时竞价的广告系统应用于多项互联网业务中，它作为互联网服务的主要利润来源与企业商业模式根基，在互联网服务业有着至关重要的作用。

实时竞价广告等智能系统的建模与优化主要有三个方面的基础模块：对于用户反馈与竞价策略的建模与分析，针对建模结果的预测与评估，实际在线竞价过程中的决策与使用。其中，建模、预测与决策这三个方面相辅相成。具体来说，建模与预测为决策提供了支持，而决策的结果与反馈又为前两者提供了优化。在人工智能应用于社会生产预测与决策下，实时竞价广告中的建模与决策优化的方法能够为其它通用智能系统带来较

多的启发，愈加受到重视<sup>[7, 9, 20]</sup>。

对于生产者及服务提供商而言，实时竞价广告中的决策模块作为实际应用于生产与平台服务的直接执行部分，对于生产质量以及服务的最终效果有着决定性影响。从更加通用的角度，这些服务与产品应用的网络服务平台包括广告投放平台的实时竞价系统<sup>[9]</sup>、智能推荐系统<sup>[19, 21]</sup>、工业生产领域的库存与流水线管理系统等，它们的决策行为包括连续实时的智能广告投放、商品的个性化推荐以及库存与生产的自动化调度等。从这些具体的社会应用中我们可以看到，决策优化与建模直接体现了其中人工智能系统的优化策略与效果，对于生产与服务流程具有根本性的影响。

对于产品与终端用户而言，决策优化的最终结果直接作用于产品与用户。在线上推荐系统与广告投放平台中，决策的操作目标是商品、广告创意及用户<sup>[22]</sup>。决策的基本原则是在合适的时间将合适的内容推荐、投放给合适的用户，并最优化整体平台效益，而效益的主要来源是用户的正反馈<sup>[16]</sup>行为。决策所依赖的模型是基于用户与上下文信息的建模与预估模型。所以，建模与决策结果直接影响了用户体验，而用户行为又对机器学习模型所处平台，例如实时竞价广告系统的整体效益有着决定性影响<sup>[23]</sup>。

对于研究者而言，机器学习相关研究经过几十年的研究，从最初的专家系统、神经网络、统计模型的建模方法，到模型应用于预测乃至决策，智能系统经历了长足的发展。现阶段，随着智能系统在生产中的大量应用，智能系统的研究方向与研究目标也逐渐转移到建模支持的决策方面，特别是在实时决策优化<sup>[9, 20, 23]</sup>方面。智能系统中的决策优化具有实时性、带限制性、反馈稀疏性等特点，所以衍生出了多种解决方案的思路。

综上所述，实时竞价广告中的建模与决策优化作为机器学习方法应用于工业应用场景最为前沿的领域，其包含的建模思想与优化方法为所在行业的及相关领域的任务提供了较多的启发，在更加通用的层面为机器学习用于人工智能系统进行预测、决策优化提供了帮助。

## 1.4 章节安排

本文所涉及内容的相关工作将在下面各章中按照各章主题分别进行介绍。文章后续章节安排如下。我们将在第二章介绍实时竞价广告的竞价机制，并介绍基于深度网络方法的生存分析方法用于建模市场竞争及市场价格，预估市场价格分布。在第三章，我们将介绍实时竞价广告中的效益预估方法，并详细介绍传统方法及其缺陷，针对性地提出直接优化广告效益与广告主收益的全新建模方案，以及对应的优化方法。第四章将聚焦在竞价优化方面，我们提出基于统一优化目标的协同优化框架，直接优化广告效益、成本预估与竞价策略，同时提供了基于广告投放效益目标的最优竞价函数以及在多智体的情况下的博弈分析。第五章我们针对线上系统中的用户建模，提出一种面向超长用户行

为序列的用户建模方案，更好地支持用户序列化行为的建模与预估。在第六章，我们针对用户转化归因问题，提出基于深度学习（**Deep Learning**）方法的多触达归因模型，并提供了基于历史广告数据的离线归因评测框架，极大地推动了计算广告中的转化归因分析研究进展。最后我们在第七章，我们针对全文研究进行总结，并展望未来工作。

## 第二章 市场建模与成本预估

如今,实时竞价广告已经占据了在线广告市场的绝大部分份额<sup>[13]</sup>。其中,对于市场竞争的建模正在越来越多地引起人们的关注。实时竞价广告一般采用“二价拍卖”(Second Price Auction)交易机制,我们将在章节2.1进行具体介绍。同时我们将市场竞争的建模定义为“竞价形势预估”(Bid Landscape Forecasting),目标是在二价交易机制下为广告主估计每一次交易市场价格分布。考虑到二价交易机制的特殊性及其导致的真实数据“删失”(censorship)问题,一般人们采用医学研究中的生存分析(Survival Analysis)方法针对广告位在单次交易中的市场价格进行预估,这部分我们将在章节2.2进行介绍。而传统方法存在着建模粒度过粗、预先对分布形式进行形式化假设以至于缺乏泛化性等问题,章节2.3将详细介绍相关工作及分析其优缺点。针对已有方法的问题,我们创新地提出了基于循环神经网络(Recurrent Neural Network)的新型深度学习方法,不仅在已有的数据集上超越了现有方法,更是在建模准确性、灵活性等方面具备独到的优势。这部分将在章节2.4进行介绍。

### 2.1 二价交易机制

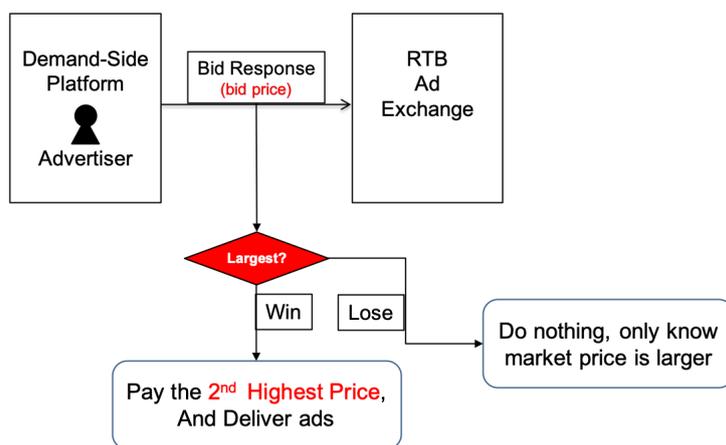


图 2-1 二价交易机制示意图

Fig 2-1 Second price auction mechanism

图2-1展示了广告主在参与二价交易时的详细交易细节。在每一次交易过程中,当所有广告主或需求方平台提出竞拍价格给广告交易所之后,广告交易所将会决定交易的赢家。按照二价交易机制的规定<sup>[7]</sup>:

- 市场上出价最高的广告主赢得本次交易，赢家将支付市场上的第二高价，即其竞争对手所出的最高竞拍价，我们称之为“市场价格” (Market Price) 或“赢得价格” (Winning Price)<sup>[24, 25]</sup>，下文一般使用市场价格进行统一表述。
- 所有未赢得交易的广告主无法获知本次交易的真实市场价格，只能获知关于市场价格的下限，即真实市场价格大于其当时的竞拍价格。

赢得交易的广告主后续将自己的广告内容展示给特定的用户，并能跟踪后续用户行为以记录广告活动中用户的反馈。

## 2.2 竞价形势预估

对于广告主来说，准确地预测市场价格是一个重要但是充满挑战的任务。一般情况下，单独为每一位参与竞价的其它广告主建模其出价策略显得不切实际<sup>[26]</sup>，因为同一时间参加实时竞价的广告主有成百上千位。另外，广告主只有预估了市场价格的分布之后，他才能够预测其任意出价情况下的胜率 (winning probability)，那么就可以估计广告成本的期望，这样才能更好地支持后续的出价决策<sup>[9]</sup>，而关于广告成本期望的计算及其与实时竞价策略优化的关系将在第三章和第四章进行介绍。市场价格分布建模与胜率预估将在本章第2.4节进行详细介绍。举例来说，有的相关文献<sup>[27]</sup> 采用了一个动态出价策略，即在预估的市场价格基础上进行微调竞拍价格；而我们将在第四章介绍一种方法利用预估的市场价格分布辅助优化出价策略函数。因此，将市场价建模为一个随机变量<sup>[25, 28]</sup>，并为每一个交易预测其市场价格在整个价格区间上的分布是一种行之有效的方案，我们称之为“竞价形势预估”，如图2-2所示。

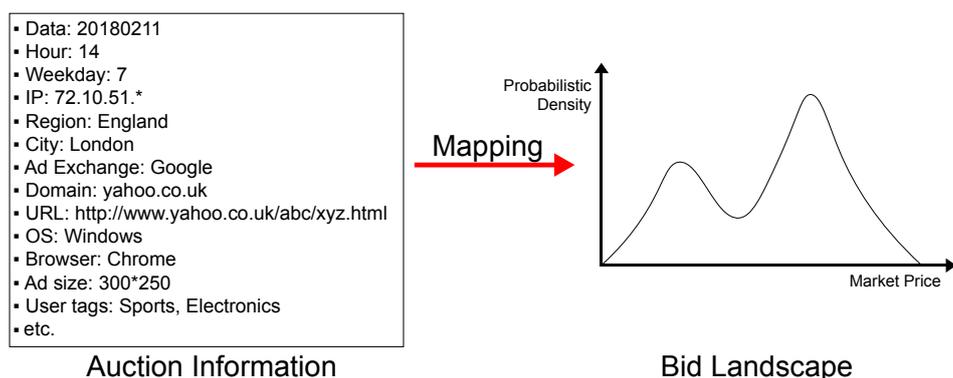


图 2-2 竞价形势预估任务

Fig 2-2 Task of bid landscape forecasting

竞价形势预估的难点主要在于三方面：

1. 无真值 (Ground Truth) 数据: 没有数据记录真实的市场价格在不同价格区间的概率;
2. 数据稀疏: 广告主的历史竞价记录数据中, 胜率一般维持在 10% 不到, 这就导致了历史数据中能够获知真实市场价格的样本较少, 直接利用回归方法对真实市场价格数据进行建模存在一定的偏差;
3. 删失问题: 对于单一广告主来说, 绝大部分交易数据无法知晓真实市场价格, 即数据删失; 而这部分交易的真实市场价格均大于每一笔交易当时该广告主的竞拍价格。这便是“幸存偏差” (Survivorship)。我们接下来具体介绍这一问题。

**幸存偏差与数据删失.** 根据上一章节关于二价交易机制的描述我们可以知道, 只有实时交易的赢家, 即出价最高的广告主才能获知交易的真实市场价格, 即该广告主获得该次广告展示机会所支付的价格; 而其他广告主无法获知真实市场价格, 仅可知晓市场价格大于其竞拍价格。所以对于大部分广告主来说, 或者对于单一广告主的大部分未赢下的历史广告交易样本, 其真实市场价格是未知的, 一般称之为数据删失; 而这部分交易的真实市场价格均大于每一笔交易当时该广告主的竞拍价格。这便是“幸存偏差” (Survivorship)。

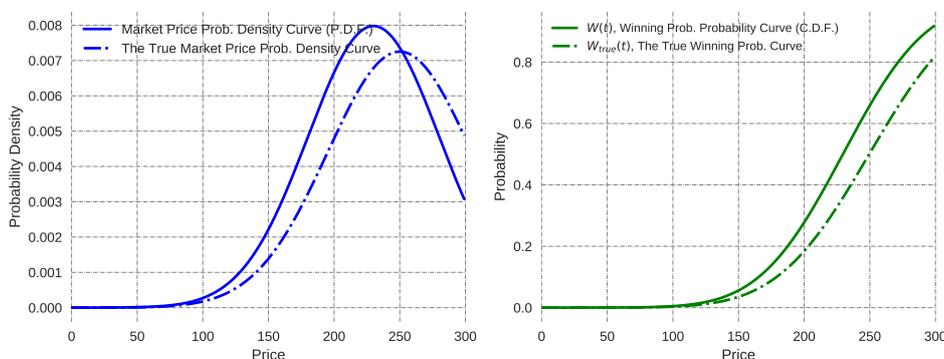


图 2-3 数据删失导致的幸存偏差图示

Fig 2-3 Illustration of censorship and survivorship

图2-3展示了由于数据删失导致的幸存偏差在市场价格分布预估与竞拍胜率预测时所造成的影响。从图中我们可以看出, 真实的市场价格分布和基于删失数据所建模得出的分布存在巨大的偏差 (bias)。有偏差的市场价格概率密度函数相较于真实概率密度函数存在着左偏的误差, 因为大量的右偏删失 (right-censored) 数据被忽略了。这些右偏删失数据的共同特典是真实值 (真实市场价格) 大于观察值 (广告主的竞拍价格)。

## 2.3 相关工作

### 2.3.1 竞价形势预估

一般研究者将竞价形势预估建模成概率分布估计问题<sup>[9, 29]</sup>。从建模的角度来说, 竞价形势预估在最近的研究历史上分为两个阶段。最初, 一些研究者提出了基于启发式的函数假设来建模市场价格分布。例如在相关文献<sup>[9, 16, 23]</sup>中, 研究者们基于竞价历史的观察结果提出了一些胜率估计函数形式的具体假设。在更近的一些研究中, 一些已被广泛应用的概率分布函数被借鉴使用在了市场价格建模的问题中。研究者在文献<sup>[24]</sup>中首次利用了 **Log-normal** 分布来建模真实市场价格的似然函数; 文献<sup>[25]</sup>利用高斯分布的假设去拟合每一次交易的真实市场价格。最近, **Gamma** 分布也被用于市场价格建模<sup>[30]</sup>。这些基于具体分布假设的方法有一个共同的问题, 即这些基于历史竞价数据的经验假设在实际的应用场景中可能并不适用, 因为实际的真实市场价格分布可能非常复杂, 如图2-4所示, 基于已有数学分布假设的方法可能缺乏泛化性 (**generalization**) 以至于无法较好建模真实的真实市场价格<sup>[29]</sup>。

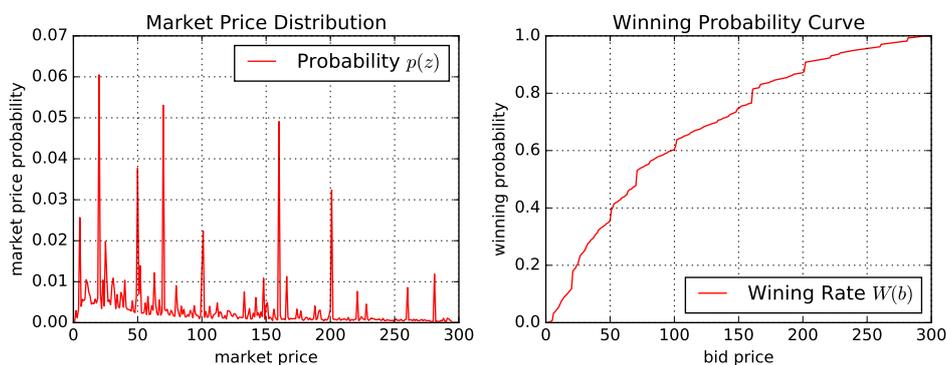


图 2-4 统计得到的真实市场价格分布与对应的胜率函数曲线

Fig 2-4 An illustration of real market price distribution and the corresponding winning probability curve

从预估 (**forecasting**) 的角度来说, 竞价形势预估的目标是为任意的交易估计其真实市场价格分布。研究者在文献<sup>[24]</sup>中利用基于模板的方法预估交易的真实市场价格分布。另一文献<sup>[29]</sup>提出了基于决策树的分类模型, 将相似的交易样本归类为同类, 并为每一类别的所有样本构建统一的非参数化 (**non-parametric**) 模型来统计市场价格分布。这两种方法都是粗粒度的建模方法, 大量相似的样本共享同一个分布建模结果, 在真实环境中难以较好地为每一个单独的交易样本进行预估。另一位研究者在文献<sup>[25]</sup>中提出了一种线性回归模型, 利用交易的特征信息直接预估其市场价格。但是上述方法都只建模了线性特征, 难以利用真实世界中高阶特征组合<sup>[15]</sup>。最近, 在文献<sup>[28]</sup>中, 研究者提

出了一种基于深度学习 (Deep Learning) 的方法用于市场价格预估。但是这类方法同样基于已有的特定数学概率模型对市场价格分布作出了假设, 亦有上文所述的缺乏泛化性的问题。同时, 这个方法直接预测每一笔交易的真实市场价格, 而非直接建模市场价格分布, 缺乏对于市场价格“不确定性”的刻画, 使得在后续实验分析中获得了较差的性能。

为了解决上述问题, 我们提出了细粒度的竞价形势预估方法, 为每一笔交易样本, 即每次单独的广告展示交易, 预估其市场价格的分布。我们利用深度神经网络提取低阶与高阶非线性的特征组合, 亦没有假设任何特定的分布函数, 使得我们的模型在建模复杂的真实市场价格分布时具备较强的灵活性, 同时获得了显著的性能提升,

### 2.3.2 删失问题处理

如章节2.2所述, 删失问题是竞价形势预估的一个重要挑战。在在线广告研究领域, 研究者一般利用基于生存分析方法的模型来一定程度上解决删失数据问题。生存分析是医学研究领域的一种典型方法, 是根据试验或调查得到的数据对生物或人的生存时间进行分析和推断, 研究生存时间和结局事件与众多影响因素间关系及其程度大小的方法<sup>1</sup>。在实时竞价广告相关工作<sup>[25, 29, 31]</sup>中, 研究者一般将“价格”看作是生存分析中的“生存时间”, 将“赢得竞价”看作是“结局事件”, 将生存分析方法间接用于市场价格(结局事件发生时间)建模。

一些文献<sup>[25, 28]</sup>利用删失回归 (Censored Regression) 方法来缓解数据偏差问题。但是他们针对真实市场价格分布采用了较强的概率分布假设, 例如高斯分布。正如前文所述, 这些分布假设在真实问题中缺乏较好的泛化性。文献<sup>[26]</sup>首次在实时竞价广告领域使用 Kaplan-Meier 生存分析模型<sup>[32]</sup>来建模市场价格分布。Kaplan-Meier 方法是一种利用统计手段来处理右偏删失数据的生存分析模型, 最早在 1958 年由 Kaplan 和 Meier 共同提出。其它研究者也在一些研究论文<sup>[29, 31]</sup>中利用这种非参数化的统计方法来建模市场价格分布与对应的竞拍胜率函数。然而 Kaplan-Meier 方法仅仅通过数学统计在一定数量的样本群体中预估概率分布, 是一种粗粒度的统计学方法, 无法针对单个样本给出个性化的预估方案。

另一种生存分析方法是 Cox 比例风险回归模型<sup>[33, 34]</sup> (Cox Proportional Hazard Model), 简称 Cox 模型。这种方法针对单位时间风险进行建模, 利用一个基本风险函数 (base hazard function) 来预测风险 (赢得竞价) 的概率, 然后通过一个基于样本特征的微调系数来调整最终风险概率。其中, 基本风险函数一般采用一些特定的数学概率模型, 例如 Weibull 分布, Log-normal 分布等。这类方法的问题同样很明显, 即针对市场价格分布

<sup>1</sup><https://zh.wikipedia.org/zh-hans/生存分析>

有确定的概率函数形态假设，在实际场景中缺乏泛化性。

另外，上述这些相关工作均没有考虑样本特征在时间维度上的序列模式 (sequential patterns)。而序列模式建模则是我们提出的新的生存分析模型的优点之一。

### 2.3.3 循环神经网络

最近，深度学习 (Deep Learning)，也就是深度神经网络 (Deep Neural Network)，由于其强大的建模能力与特征抽取效果，正在各研究领域发挥重要的作用，包括计算机视觉<sup>[35]</sup>、语音识别<sup>[36]</sup>和自然语言处理<sup>[37, 38]</sup>等。其中，循环神经网络 (Recurrent Neural Network, RNN)及其变种 LSTM<sup>[37]</sup> (Long Short-term Memory)利用自回归 (auto-regressive) 的特性，在刻画序列模式时具有显著的优势。在本章，我们利用循环神经网络的自回归特性建模条件概率，并提出一种基于深度学习的新型生存分析模型，在实验中极大地提高了预测精度，更为准确地建模了市场价格的分布，为细粒度概率分布预测提供了新的思路。

## 2.4 基于深度生存分析的方法

在本节，我们将详细介绍一种新的基于深度生存分析的方法，我们称之为深度形势预估 (Deep Landscape Forecasting)，简称 DLF 模型。章节2.4.1将对竞价形式预估进行形式化定义。之后我们将在章节2.4.2详细介绍我们的 DLF 模型。实验设置与结果分析将在章节2.5介绍。

### 2.4.1 问题定义

在实时竞价广告的场景下，如图2-1所示，广告主需要在收到每次竞拍的竞价请求后，返回竞拍价格  $b$ 。竞拍特征  $\mathbf{x}$  包含三方面的信息，用户端信息 (位置、浏览器信息等)、发行方信息 (网页网址、广告位大小等) 以及具体的待展示的广告内容信息。广告主竞价策略的主要目标是在考虑自身预算的情况下尽可能最大化效益，关于竞价策略的部分内容将在第四进行介绍。我们将市场价格，即赢下交易后所支付的市场第二高价，用变量  $z$  表示，其概率密度函数 (probabilistic density function, P.D.F.) 是  $p(z), z > 0$ 。

注意到，给定任意竞拍价格  $b$ ，我们能通过市场价格的概率密度函数得到竞拍胜率  $W(b)$  如下

$$W(b) \doteq \Pr(z < b) = \int_0^b p(z) dz, \quad (2-1)$$

按照定义，胜率就是竞拍价格大于市场价格的概率。那么，广告竞拍的输率即胜率的逆

事件

$$S(b) \doteq \Pr(z \geq b) = 1 - W(b) = \int_b^{\infty} p(z)dz. \quad (2-2)$$

竞拍历史数据由一系列三元组  $\{(\mathbf{x}, b, z)\}$  构成，在每个三元组中， $\mathbf{x}$  是竞拍请求特征信息， $b$  是当时交易中广告主的竞拍价格， $z$  是交易的市场价格。注意，只有当广告主当时赢下了竞拍， $z$  的值才能被知晓并记录；当广告主没有赢下该竞价交易时， $z$  的值位置，我们将其标记为 **Null**。

竞价形势预估的主要任务是，给定竞拍请求特征信息  $\mathbf{x}$ ，广告主需要预测市场价格的概率密度函数  $p(z|\mathbf{x})$ 。形式上来说，我们寻求的预估函数  $T$  是将竞价信息“映射”为概率分布的函数

$$p(z|\mathbf{x}) = T(\mathbf{x}). \quad (2-3)$$

## 2.4.2 模型介绍

本章节将具体介绍本文提出的深度形势预估模型 **DLF**。

### 2.4.2.1 离散价格空间

不失一般性，首先我们将连续 (**continuous**) 价格空间转换为离散 (**discrete**) 价格空间。因为在实时竞价广告场景下，所有的价格信息都是离散的整数，所以基于离散空间的讨论具有现实意义，并已被相关文献<sup>[29, 31]</sup> 广泛采用。

我们将连续的价格切分为  $L$  个离散区间，并定义一系列离散价格  $0 < b_1 < b_2 < \dots < b_L$ 。同时我们将离散价格区间标记为  $V_l = (b_l, b_{l+1}]$ ,  $l = 1, 2, \dots, L$ ，其中  $b_0 = 0$ ， $b_l$  是竞拍交易样本中的竞拍出价，同时  $b_{l+1} - b_l = 1$ 。

定义在离散价格空间上的胜率函数和输率函数分别为

$$\begin{aligned} W(b_l) &\doteq \Pr(z < b_l) = \sum_{j < l} \Pr(z \in V_j), \\ S(b_l) &\doteq \Pr(b_l \leq z) = \sum_{j \geq l} \Pr(z \in V_j), \end{aligned} \quad (2-4)$$

这两个函数的输入则是广告主出价  $b_l$ 。那么定义在离散价格空间的市场价格分布列是

$$\begin{aligned} p_l &= \Pr(z \in V_l) = W(b_{l+1}) - W(b_l) \\ &= [1 - S(b_{l+1})] - [1 - S(b_l)] \\ &= S(b_l) - S(b_{l+1}). \end{aligned} \quad (2-5)$$

接下来, 我们定义给定竞拍价格  $b_l$  时, 广告主的条件胜率 (conditional winning probability) 函数如下

$$h_l = \Pr(z \in V_l | z \geq b_{l-1}) = \frac{\Pr(z \in V_l)}{\Pr(z \geq b_{l-1})} = \frac{p_l}{S(b_{l-1})}, \quad (2-6)$$

这个条件概率代表了广告主出更低价格没能赢下竞拍的条件, 恰好在出价  $b_l$  时赢下交易的概率。这也是市场价格大于更小的价格 (条件) 但恰好落在第  $l$  个价格区间  $V_l = (b_l, b_{l+1}]$  (事件) 的条件概率。

#### 2.4.2.2 循环神经网络模型

接下来我们具体阐述基于循环神经网络的新型生存分析模型。

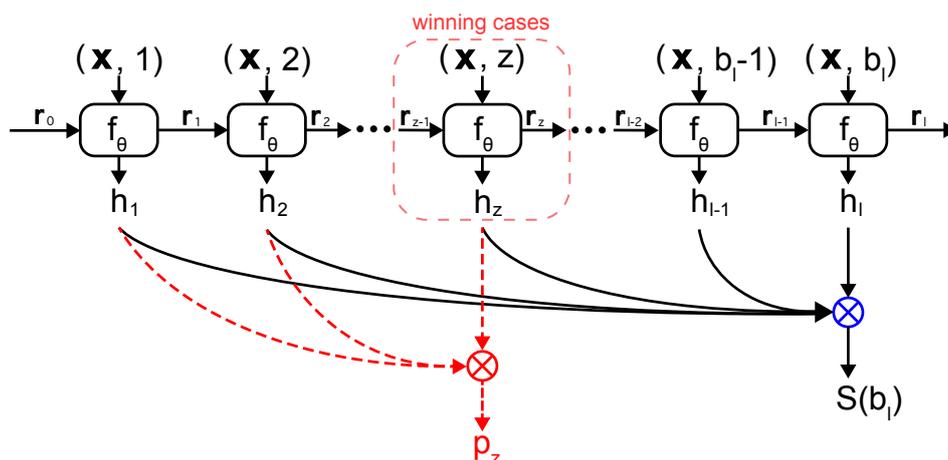


图 2-5 深度形势预估模型的框架图

Fig 2-5 Detailed illustration of DLF model

模型的整体结构如图2-5所示。给定任意竞拍特征数据  $\mathbf{x}$ , DLF 模型为任意第  $l$  个价格区间建模条件胜率函数  $h_l^i$  如下

$$\begin{aligned} h_l^i &= \Pr(z \in V_l | z \geq b_{l-1}, \mathbf{x}^i; \theta) \\ &= f_{\theta}(\mathbf{x}^i, b_l | \mathbf{r}_{l-1}), \end{aligned} \quad (2-7)$$

这里  $f_{\theta}$  是以  $\theta$  为参数的循环神经网络, 他每个时刻的其中一部分输入是  $(\mathbf{x}^i, b_l)$ , 输出  $h_l^i$  和隐状态向量 (hidden state vector)  $\mathbf{r}_{l-1}$ 。另一部分输入  $\mathbf{r}_{l-1}$  是上一个时刻神经网络输出的隐状态向量。在我们的具体实现中, 我们使用了 LSTM<sup>[37]</sup> 来实现具体的神经网络单元, 具体实现将在附录D中详细介绍。

通过公式 (2-4), (2-6) 和 (2-7), 我们能够比较容易地得到给定任意竞拍价  $b$  的情况下胜率函数  $W(b)$  与输率函数  $S(b)$

$$\begin{aligned}
S(b|\mathbf{x}^i; \boldsymbol{\theta}) &= \Pr(b \leq z|\mathbf{x}^i; \boldsymbol{\theta}) = \Pr(z \notin V_1, z \notin V_2, \dots, z \notin V_l|\mathbf{x}^i; \boldsymbol{\theta}) \\
&= \Pr(z \notin V_1|\mathbf{x}^i; \boldsymbol{\theta}) \cdot \Pr(z \notin V_2|z \notin V_1, \mathbf{x}^i; \boldsymbol{\theta}) \cdots \\
&\quad \cdot \Pr(z \notin V_l|z \notin V_1, \dots, z \notin V_{l-1}, \mathbf{x}^i; \boldsymbol{\theta}) \\
&= \prod_{l:l \leq l_b} [1 - \Pr(z \in V_l | z \geq b_l, \mathbf{x}^i; \boldsymbol{\theta})] = \prod_{l:l \leq l_b} (1 - h_l^i), \\
W(b|\mathbf{x}^i; \boldsymbol{\theta}) &= \Pr(b > z|\mathbf{x}^i; \boldsymbol{\theta}) = 1 - S(b|\mathbf{x}^i; \boldsymbol{\theta}) = 1 - \prod_{l:l \leq l_b} (1 - h_l^i),
\end{aligned} \tag{2-8}$$

这里的  $l_b$  是给定竞拍价格  $b$  所在的价格区间。在这里的公式推导过程中, 我们使用了概率链式法则来计算输率  $S(b)$ 。具体方法是将每一个条件概率连乘获得最后的联合概率。

之后, 考虑公式 (2-5) 和 (2-6), 任意竞拍交易的市场价格的分布列函数可以通过相似推导获得

$$p_l^i = \Pr(z^i \in V_l|\mathbf{x}^i; \boldsymbol{\theta}) = h_{l_i}^i \cdot S(b = b_{l_i} - 1|\mathbf{x}^i; \boldsymbol{\theta}) = h_{l_i}^i \prod_{l:l < l^i} (1 - h_l^i). \tag{2-9}$$

### 2.4.2.3 损失函数设计

在本小节, 我们详细介绍这种基于深度学习的生存分析模型的损失函数设计。需要注意的是, 数据中没有任何关于市场价格概率分布函数或分布列的真值信息, 所以我们需要通过最大似然估计的方法来监督模型的优化。具体来说, 我们采用了两种损失函数。

第一种损失函数基于分布列  $p_l^i$ , 目标是在赢下的竞价历史数据中减小市场价格值的负对数似然, 即

$$\begin{aligned}
L_1 &= -\log \prod_{(\mathbf{x}^i, z^i) \in \mathbb{D}_{\text{win}}} \Pr(z^i \in V_{l^i}|\mathbf{x}^i; \boldsymbol{\theta}) = -\log \prod_{(\mathbf{x}^i, z^i) \in \mathbb{D}_{\text{win}}} p_{l^i}^i \\
&= -\log \prod_{(\mathbf{x}^i, z^i) \in \mathbb{D}_{\text{win}}} h_{l^i}^i \prod_{l:l < l^i} (1 - h_l^i) \\
&= - \sum_{(\mathbf{x}^i, z^i) \in \mathbb{D}_{\text{win}}} \left[ \log h_{l^i}^i + \sum_{l:l < l^i} \log(1 - h_l^i) \right],
\end{aligned} \tag{2-10}$$

这里  $l^i$  是给定交易特征  $\mathbf{x}^i$  时, 市场价格  $z^i \in V_{l^i}$  所在的价格区间标记。

第二种损失函数基于累积分布函数 (cumulative distribution function, C.D.F.), 即胜率函数。我们在图2-6中列举了两个例子。左图是一个广告主赢得竞价交易的例子, 右

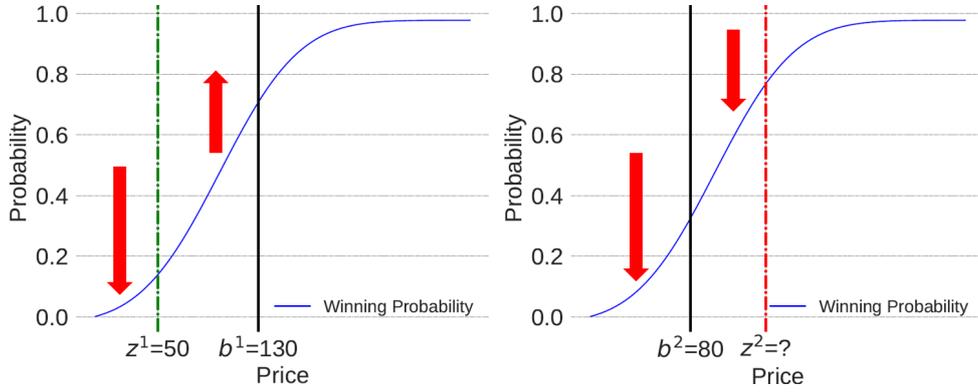


图 2-6 胜率函数曲线的两种例子

Fig 2-6 Two examples of winning curves

图则是广告主未赢得交易的例子。对于赢下的交易，我们知道竞拍价格大于市场价格，即  $b^1 \leq z^1$ ；而对于未赢的交易，我们无法知晓真实的市场价格，只能获知  $b^2 < z^2$ 。那么我们对于第二种损失函数有两方面的动机考虑。

直观上来说，对于左图赢下的交易，我们需要将胜率函数曲线在  $[0, z]$  区间向下“压”；而在  $[z, \infty)$  区间往上“提”，特别是在  $[z, t]$  价格区间的胜率函数曲线段。因此，我们设计第二种损失函数的其中一种形式为

$$\begin{aligned}
 L_{\text{win}} &= -\log \prod_{(\mathbf{x}^i, b^i) \in \mathbb{D}_{\text{win}}} \Pr(b^i > z | \mathbf{x}^i; \boldsymbol{\theta}) \\
 &= -\log \prod_{(\mathbf{x}^i, b^i) \in \mathbb{D}_{\text{win}}} W(b^i | \mathbf{x}^i; \boldsymbol{\theta}) \\
 &= -\sum_{(\mathbf{x}^i, b^i) \in \mathbb{D}_{\text{win}}} \log \left[ 1 - \prod_{l:l \leq i} (1 - h_l^i) \right].
 \end{aligned} \tag{2-11}$$

对于右图未赢的交易，我们需要胜率函数曲线往下“压”即可，因为我们无法确知真实市场价格  $z$ ，仅仅知道  $z > t$ 。那么我们设计了第二种损失函数的另一形式为

$$\begin{aligned}
 L_{\text{lose}} &= -\log \prod_{(\mathbf{x}^i, b^i) \in \mathbb{D}_{\text{lose}}} \Pr(z \geq b^i | \mathbf{x}^i; \boldsymbol{\theta}) \\
 &= -\log \prod_{(\mathbf{x}^i, b^i) \in \mathbb{D}_{\text{lose}}} S(b^i | \mathbf{x}^i; \boldsymbol{\theta}) \\
 &= -\sum_{(\mathbf{x}^i, b^i) \in \mathbb{D}_{\text{lose}}} \sum_{l:l \leq i} \log(1 - h_l^i).
 \end{aligned} \tag{2-12}$$

对于任意的交易样本，我们可以得到一个关于“赢得竞价”的指示函数

$$w^i = \begin{cases} 1, & \text{若 } b^i > z^i, \\ 0, & \text{否则 } b^i \leq z^i. \end{cases} \quad (2-13)$$

对于赢得的交易样本  $(\mathbf{x}^i, z^i, b^i)$  来说， $z^i$  已知未删失，所以  $w^i = 1$ ；而对于未赢得的交易样本来说， $z^i$  删失未知具体值，仅仅知道  $z^i \geq b^i$ ，故  $w^i = 0$ 。

结合第二种损失函数的两种形式 (2-11) 和 (2-12) 来看， $L_{\text{win}}$  和  $L_{\text{lose}}$  其实是基于累积分布函数在做胜率预测。我们由此得到了第二种损失函数的综合形式

$$\begin{aligned} L_2 &= L_{\text{win}} + L_{\text{lose}} \\ &= -\log \prod_{(\mathbf{x}^i, b^i) \in \mathbb{D}_{\text{win}}} \Pr(b^i > z | \mathbf{x}^i; \boldsymbol{\theta}) - \log \prod_{(\mathbf{x}^i, b^i) \in \mathbb{D}_{\text{lose}}} \Pr(z \geq b^i | \mathbf{x}^i; \boldsymbol{\theta}) \\ &= -\log \prod_{(\mathbf{x}^i, b^i) \in \mathbb{D}} [W(b^i | \mathbf{x}^i; \boldsymbol{\theta})]^{w^i} \cdot [1 - W(b^i | \mathbf{x}^i; \boldsymbol{\theta})]^{1-w^i} \\ &= -\sum_{(\mathbf{x}^i, b^i) \in \mathbb{D}} \{w^i \cdot \log W(b^i | \mathbf{x}^i; \boldsymbol{\theta}) + (1 - w^i) \log [1 - W(b^i | \mathbf{x}^i; \boldsymbol{\theta})]\}, \end{aligned} \quad (2-14)$$

这是关于广告主竞拍价格的胜率预估结果的交叉熵，能够基于所有的交易样本  $\mathbb{D} = \mathbb{D}_{\text{win}} \cup \mathbb{D}_{\text{lose}}$  进行学习而训练（注意第一种损失函数  $L_1$  仅仅只能基于已赢得的交易样本  $\mathbb{D}_{\text{win}}$  进行训练和学习，如公式2-10所示）。第二种损失函数针对删失数据进行了建模，能够较大程度地缓解右偏删失导致的预测偏差，这点将在后面实验部分进行印证。

另外，在传统的生存分析方法<sup>[33, 39]</sup>以及竞价形势预估相关工作<sup>[25, 30]</sup>中，研究者一般只用了  $L_1$  损失函数和  $L_{\text{lose}}$ 。而在我们的方法中，我们创新性地将赢得交易样本与未赢得交易样本进行联合建模，能够更好地基于累积分布函数进行学习。

综上，我们最终的损失函数是

$$\arg \min_{\boldsymbol{\theta}} \alpha L_1 + (1 - \alpha) L_2, \quad (2-15)$$

其中  $\alpha$  控制了两种损失函数计算所得的学习梯度数量级保持一致。

#### 2.4.2.4 模型效能

本章节将分析 DLF 模型的计算复杂性。正如公式 (2-7) 所列，每一个循环神经网络单元  $f_{\boldsymbol{\theta}}$  输入  $(\mathbf{x}, b_l, \mathbf{r}_{l-1})$  并输出标量  $h_l$  和隐向量  $\mathbf{r}_l$  给下一时刻的网络单元。每笔交易需要考虑的最多  $L$  个价格区间，所以循环神经网络单元需要最多计算  $L$  次。我们设定循环神经单元的计算复杂度为  $O(C)$ ，具体的计算复杂度与神经网络单元的具体设计有

关<sup>[40]</sup>，并且可以被图形处理单元 (Graphical Processing Unit, GPU) 并行加速。在我们的模型中，为了获得市场价格分布列  $p_z$  和胜率  $W$ ，需要将所有的条件胜率输出  $h_l$  (或者其逆事件概率  $(1 - h_l)$ ) 连乘，而这里的计算复杂度则是  $O(L)$ 。所以总体模型时间复杂度是  $O(CL) + O(L) = O(CL)$ ，与原始循环神经网络的复杂度一致。

在许多相关工作中，循环神经网络已被广泛应用于推荐系统<sup>[41]</sup>、在线广告平台<sup>[42, 43]</sup>和在线机器翻译系统<sup>[44]</sup>。在我们的实验中，基于一般被广泛推荐的系统设定，我们测试 DLF 模型的单次推理 (inference) 时间约为 22 毫秒，能够满足前文所述的实时竞价广告流程的总体耗时要求 (100 毫秒<sup>[7]</sup>)。

## 2.5 实验分析

在接下来的章节中，我们将详细介绍实验细节，包括章节 2.5.1 对于数据集与实验总体流程的介绍、章节 2.5.2 对于评估指标与对比模型的介绍以及章节 2.5.3 详细分析的实验结果。

### 2.5.1 数据集与实验流程

**数据集.** 实验使用了两个实时竞价广告真实数据集。iPinYouRTB 数据集最早发布在相关文献<sup>[45]</sup>中。它包含了 6,470 万条竞价历史样本，1950 万条广告展示 (广告主只有赢下交易后才能进行广告展示)，14,790 个点击以及总计 16,000 元人民币的广告投放开销。这个数据集由 9 个广告活动 (campaigns) 子数据集构成，收集自这些广告活动在 2013 年总共为期 10 天的投放。每一条广告样本总共包含 16 种属性信息，包括时间信息、用户信息、区域信息、广告位信息等。我们将前 7 天的数据设定为训练数据，最后 3 天的数据设定为测试数据。另外一个数据集 YOYI 最早由另一个文献<sup>[6]</sup>发布。它包含 40,200 万条广告展示样本，500,000 个点击以及总计 428,000 元人民币的广告投放。这个数据集包含 2016 年 1 月总计 8 天的实时竞价广告投放数据。更多关于这两个数据集的信息可以参考相关文献<sup>[16, 45]</sup>。

**数据预处理.** 为了模拟真实竞价环境，并突出展示本文所提出的新型生存分析方法的优越性，我们将原始竞价历史数据作为模拟竞价的数据源，并使用真实出价策略<sup>1</sup> (truthful bidding strategy) 进行竞价模拟，来生成赢下的竞价历史数据集  $\mathbb{D}_{\text{win}}$  和未赢的竞价历史数据集  $\mathbb{D}_{\text{lose}}$ ，对于前者的每一条竞价数据  $(\mathbf{x}_{\text{win}}, b_{\text{win}}, z_{\text{win}}) \in \mathbb{D}_{\text{win}}$ ，真实市场价格  $z_{\text{win}} < b_{\text{win}}$  被广告主所知晓；对于后者的每一条竞价数据  $(\mathbf{x}_{\text{lose}}, b_{\text{lose}}, z_{\text{lose}}) \in \mathbb{D}_{\text{lose}}$ ，真实市场价格未知但广告主知道  $z_{\text{lose}} \geq b_{\text{lose}}$ 。这个模拟符合现实场景中实时竞价广告的必要条件，也保

<sup>1</sup>更多关于“真实出价策略”的介绍将在第三章和第四章进行详细介绍。

表 2-1 竞价形势预估离线实验数据集统计, WR: 平均胜率; AMP: 平均市场价

Table 2-1 The statistics of the datasets. WR: winning rate; AMP: averaged market price.

广告活动	总条数	赢下交易数	WR	AMP	AMP ( $\mathbb{D}_{\text{win}}$ )	AMP ( $\mathbb{D}_{\text{lose}}$ )
1458	3,697,694	1,116,644	0.3020	69.6696	27.4265	87.9452
2259	1,252,753	396,283	0.3163	96.7888	27.1986	128.9877
2261	1,031,479	321,931	0.3121	87.6479	18.9000	118.8396
2821	1,984,525	228,833	0.1153	93.8962	13.2118	104.4125
2997	468,500	70,747	0.1510	60.4188	7.2762	69.8711
3358	2,043,032	315,010	0.1542	95.4967	21.2540	109.0308
3386	3,393,223	819,447	0.2415	78.0327	23.8983	95.2682
3427	3,130,560	654,989	0.2092	81.9650	25.2118	96.9808
3476	2,494,208	723,847	0.2902	80.0719	31.2218	100.0453
总计	19495974	4647731	0.2384	82.0744	25.0484	99.9244
YOYI	401,617,064	202,214,191	0.5035	55.7444	24.4488	87.4842

证了本文的离线实验符合现实环境。同时, 这个模拟设置也在大量竞价形势预估的相关文献<sup>[25, 29-31]</sup> 中被广泛采用。

经过数据预处理后, 我们做了一些统计, 结果陈列于表格2-1中。从表格中我们可以发现, 赢下的交易样本平均市场价远低于未赢的交易样本的平均市场价, 可见删失样本对于预测结果的影响非常巨大。另外, 因为所有价格都是离散整数, 所以我们设定模型的最大价格区间为  $L = 300$ 。

**实验流程.** 对于每一个交易样本  $(\mathbf{x}, b, z)$ , 其对应的市场价格分布函数  $p(z|\mathbf{x})$  或分布列将由下面将介绍的每一个对比模型分别给出, 然后其对应的胜率函数  $W(b|\mathbf{x})$  与输率函数  $S(b|\mathbf{x})$  也可以由公式 (2-1)、(2-2) 和 (2-4) 计算得到。接下来, 每个模型的输出将由下面介绍的评价指标进行性能评估。

## 2.5.2 评估指标与对比模型

**评估指标.** 第一个评估指标是平均负对数似然概率 (averaged negative log probability, **ANLP**), 最早在 2016 年相关文献<sup>[29]</sup> 中提出。它直接用于评估市场价格分布预估的精确程度, 计算方法如下

$$\bar{P} = -\frac{1}{|\mathbb{D}_{\text{test}}|} \sum_{(\mathbf{x}^i, z^i) \in \mathbb{D}_{\text{test}}} \log p_z(z^i|\mathbf{x}^i), \quad (2-16)$$

其中  $p_z(z|\mathbf{x})$  是每个模型在测试数据上的预测输出结果。

第二个评估指标是 **C 指数** (concordance index, **C-index**), 它是生存分析相关工作中使用最多的一种评估指标, 用于衡量给定竞拍价格下模型预测胜率的准确率。对于竞拍价格  $b$ , 给定两个交易样本,  $d^1 = (\mathbf{x}^1, z^1)$  拥有相比  $b$  较大的市场价格  $z^1 \geq b$  而

$d^2 = (x^2, z^2)$  拥有相比  $b$  更小的市场价格  $z^2 < b$ , 那么第一个交易样本的排序必须在第二个之前, 即  $d^1 \prec d^2$ 。这个评估指标和二分类问题<sup>[15, 16]</sup>中的 ROC 曲线面积 (area under ROC curve, AUC) 指标是一致的 (当且仅当生存分析建模中只有一种事件时)。

最后, 我们针对模型输出还做了显著性测试 (significance test)。具体来说, Mann-Whitney U 测试<sup>[46]</sup> 被用于 C 指数, t 测试<sup>[47]</sup> 被用于 ANLP 指标。

**对比模型.** 实验中我们将 DLF 模型对比了 9 个基线模型, 包括生存分析方法及一些深度学习学习方法等。

- KM** 是前文所述的 Kaplan-Meier 无参数统计方法<sup>[32]</sup>。
- Lasso-Cox** 半参数化的方法<sup>[34]</sup>, 它基于 Cox 回归生存分析模型<sup>[33]</sup> 并使用  $l_1$  正则处理。
- DeepSurv** 是一种使用深度学习方法进行特征抽取的 Cox 回归分析方法<sup>[39]</sup>。
- Gamma** 是一种基于 Gamma 分布的回归方法<sup>[30]</sup>, 它假设每一笔交易的市场价格服从 Gamma 分布。
- MM** 是混合回归方法<sup>[25]</sup>, 它结合了线性回归和偏差学习。
- MTLSA** 是一种基于多任务学习的生存分析方法<sup>[48]</sup>。
- STM** 是一种生存树模型<sup>[29]</sup>, 他结合了 Kaplan-Meier 方法和决策树, 也是竞价形势预估任务中的最先进的模型 (state-of-the-art)。
- DeepHit** 是一种无分布假设的深度学习生存分析模型<sup>[49]</sup>, 它直接利用深度神经网络预测样本的市场价格分布列, 但没有像我们一样利用序列信息。
- DWPP** 是一种基于深度神经网络的市场价格预测模型<sup>[28]</sup>, 它和 Gamma 模型一样对市场价格分布做了一定的假设, 我们在实验中使用了原文里面表现最好的 Gaussian 分布。
- RNN** 是基于 DLF 模型的简化模型, 它只有第一种损失函数  $L_1$ , 没有考虑删失数据。该模型将用于实验分析中对于删失数据建模的处理。
- DLF** 是本文提出的模型, 具体介绍在章节 2.4.2 中。

### 2.5.3 实验结果与对比分析

本节将具体呈现实验结果并给出详细的分析。

**竞价形势预估性能.** 我们首先分析不同模型对于竞价形势预估性能, 即市场价格分布的预估性能。表格 2-3 列出了每个模型 ANLP 指标的详细评估结果。从表格中我们可以发现, DLF 模型在其它基线模型的基础上显著提升了 ANLP 指标性能, 包括之前的最先进模型 STM。

另外我们还能发现, 基于树模型的生存分析方法 STM 相比较其它基线模型有较大

表 2-3 平均负对数似然概率 ANLP 评估结果, 数值越低代表性能越好。(\* 代表显著性分析中  $p\text{-value} < 10^{-6}$ )

Table 2-3 Averaged negative log probability performance: the smaller, the better. (\* indicates  $p\text{-value} < 10^{-6}$  in significance test)

iPinYou	ANLP										
	KM	Lasso-Cox	DeepSurv	Gamma	MM	MTLSA	STM	DeepHit	DWPP	RNN	DLF
1458	10.532	38.608	38.652	5.956	5.788	9.791	4.761	5.510	29.204	9.506	<b>4.088*</b>
2259	14.671	28.234	29.658	6.069	7.328	10.248	5.471	5.586	39.263	9.625	<b>5.244*</b>
2261	14.665	39.129	39.390	5.986	7.020	10.261	4.818	5.442	32.805	9.417	<b>4.632*</b>
2821	19.582	43.099	43.072	7.838	7.262	9.895	5.572	5.614	40.537	23.099	<b>5.428*</b>
2997	16.203	32.849	33.052	5.999	6.702	9.167	5.083	5.470	34.940	16.639	<b>4.504*</b>
3358	19.253	44.769	44.885	6.736	7.177	9.484	5.539	5.616	40.958	13.806	<b>5.281*</b>
3386	15.973	39.781	41.943	6.488	6.141	8.834	5.228	5.549	32.550	10.743	<b>4.940*</b>
3427	16.902	41.558	41.698	6.002	6.185	9.090	5.321	5.552	33.387	9.565	<b>4.836*</b>
3476	10.507	39.551	39.518	5.710	6.022	10.240	4.537	5.554	31.609	7.891	<b>4.012*</b>
Overall	15.366	38.620	39.096	6.310	6.552	9.668	5.148	5.544	35.028	12.255	<b>4.774*</b>
YOYI	7.907	30.946	27.897	6.475	5.652	10.286	4.503	5.567	29.108	5.885	<b>4.453*</b>

的性能提升, 这可能得益于其较好的决策树分类能力。所有的生存分析方法, 即 DeepSurv、Gamma、MM、STM、MTLSA 和 DeepHit 均比 RNN 的性能好很多, 这也从侧面反映出了对于删失数据的建模将有效提升模型预测能力。虽然都是没有假设数据分布类型的深度学习方法, 我们的 DLF 模型相比 DeepHit 有显著优势, 原因是我们模型在公式 (2-7) 中对于序列模式的较好挖掘, 而 DeepHit 完全忽略了序列模式。虽然 DWPP 利用深度学习做特征抽取, 但是它的 ANLP 性能较差, 这个发现也在作者原文中获得了体现<sup>[28]</sup>。其原因可能是对于市场价格分布的形式化假设 (Gaussian 分布) 缺乏泛化性。

表 2-4 C 指数评估结果, 数值越高代表性能越好。(\* 代表显著性分析中  $p\text{-value} < 10^{-6}$ )

Table 2-4 C-index performance: the larger, the better, (\* indicates  $p\text{-value} < 10^{-6}$  in significance test)

iPinYou	C-index										
	KM	Lasso-Cox	DeepSurv	Gamma	MM	MTLSA	STM	DeepHit	DWPP	RNN	DLF
1458	0.698	0.820	0.835	0.612	0.698	0.505	0.764	0.861	0.866	0.894	<b>0.904*</b>
2259	0.685	0.775	0.791	0.584	0.685	0.505	0.768	0.785	0.729	0.791	<b>0.876*</b>
2261	0.666	0.847	0.890	0.564	0.666	0.508	0.812	0.838	0.807	0.874	<b>0.929*</b>
2821	0.677	0.741	0.714	0.563	0.678	0.507	0.790	0.810	0.746	0.737	<b>0.881*</b>
2997	0.734	0.910	0.852	0.641	0.734	0.517	0.835	0.907	0.885	0.762	<b>0.919*</b>
3358	0.704	0.866	0.896	0.601	0.706	0.542	0.811	0.888	0.744	0.819	<b>0.944*</b>
3386	0.716	0.845	0.854	0.569	0.719	0.512	0.849	0.881	0.833	0.800	<b>0.923*</b>
3427	0.724	0.830	0.845	0.586	0.742	0.508	0.798	0.873	0.796	0.804	<b>0.901*</b>
3476	0.692	0.865	0.877	0.676	0.692	0.505	0.830	0.879	0.861	0.917	<b>0.922*</b>
Overall	0.700	0.834	0.840	0.600	0.703	0.513	0.807	0.858	0.807	0.823	<b>0.911*</b>
YOYI	0.791	0.847	0.862	0.528	0.791	0.510	0.886	0.878	0.856	0.898	<b>0.924*</b>

**胜率估计性能评估.** 表格2-4罗列了 C 指数评估指标的模型性能结果, 这个指标直接衡量了基于累积概率函数的胜率预估准确性。从表格中亦可以发现, 我们的 DLF 模型再次在性能指标上获得了显著的更好表现。特别地, 我们模型相比较之前的最先进模型 STM 获得了平均 12.9% 的提升。

另外, 表格结果还显示出一些新的发现。所有基于深度学习的方法包括 DeepSurv、DeepHit、DWPP、RNN 和 DLF 相比较其他非深度学习方法均有更好的性能表现。甚至是没有建模删失数据的 RNN 模型也有相对不错的性能结果, 这也从侧面反映出我们对于序列模式建模的必要性及其带来的性能优势。MM 和 Gamma 模型表现较差, 这可能是由于这两个模型对于市场价格分布做的预先假设使得模型泛化性能降低, 不适应于真实场景的市场价格数据。

**模型收敛性能.** 为了展示我们 DLF 模型的收敛性能, 图2-7绘制了模型的学习曲线, 四

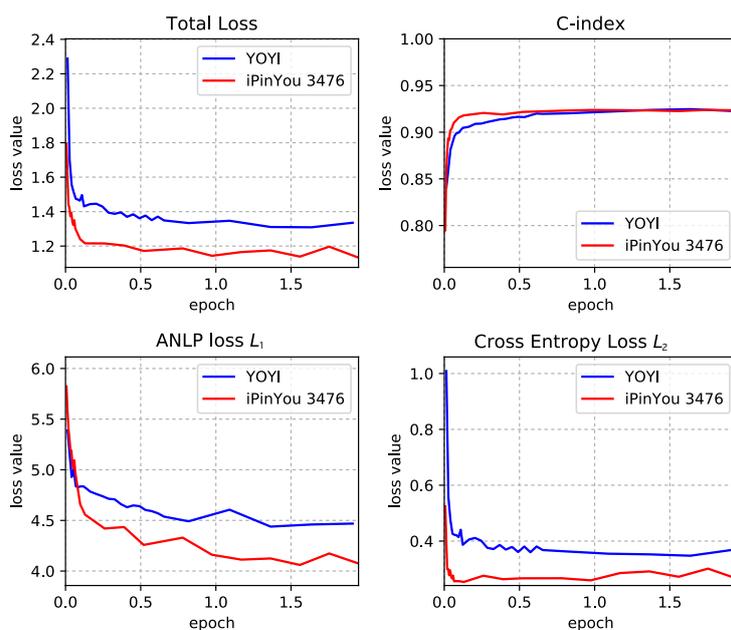


图 2-7 基于 iPinYou 数据集 3476 与 YOYI 数据集绘制的模型学习曲线

Fig 2-7 Learning curves over Campaign 3476 of iPinYou and YOYI

个图分别展示的是总损失函数 (total loss)、C-index 指标、ANLP 损失目标  $L_1$  和交叉熵损失目标  $L_2$  随着训练迭代的进行的输出值变化。这里的“epoch”指的是整个训练集的完整迭代, 我们还设置了公式 (2-15) 中的  $\alpha = 0.25$ 。从图中我们可以认为, 模型在两种损失函数的指导下收敛情况良好, 在一次完整训练集学习迭代之后较快收敛, 并且两种损失函数能协同进行监督学习。

**可视化分析.** 在图2-8中, 我们选取了一个具体的交易样本并将所有对比模型的输出进

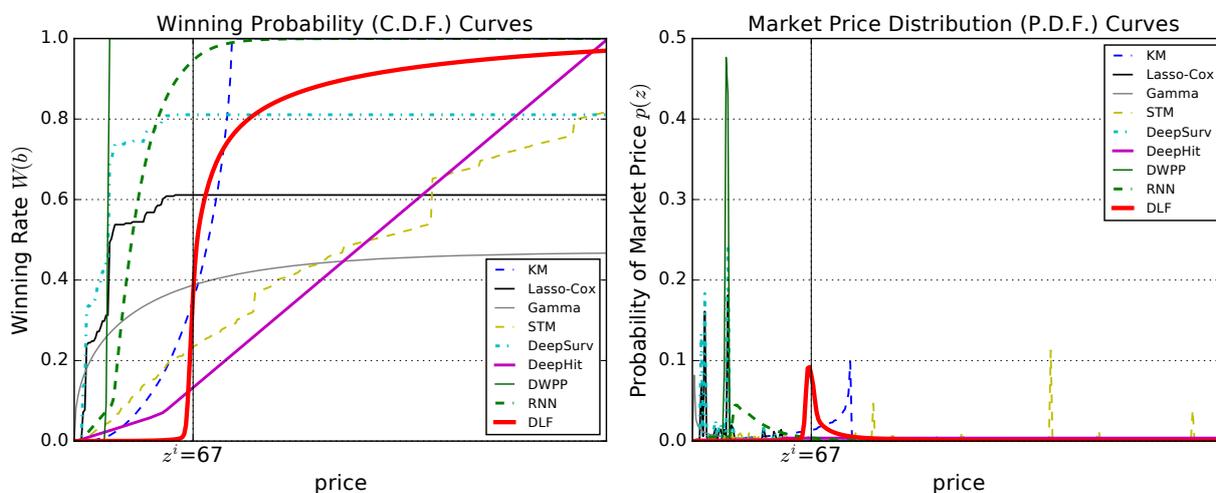


图 2-8 各模型输出结果可视化

Fig 2-8 A comprehensive visualization the forecasted results from different models

行可视化，输出包括胜率函数预估曲线  $W(b|\mathbf{x}^i)$  和市场价格分布函数或分布列  $p(z|\mathbf{x}^i)$ 。这个例子中真实的市场价格是  $z^i = 67$ 。对于竞价形势预估任务来说，表现越好的模型，其对于真实市场价格的分布估计越准确。从图中我们可以发现，我们的 DLF 模型准确地将市场价格及其附近的概率密度置予了最大权重，胜率曲线在这一点附近有一个极大的突变，而其他模型则并没有合理地将概率密度分布在整个价格空间内。这个现象反映出我们的模型在竞价形式预估任务中优秀的预估准确率。对比 RNN 和 DLF 模型我们还发现，RNN 较为“乐观”地将胜率提早估计得较高，这是因为它未能建模删失数据，删失数据中包含的信息是真实市场价格相比当时的竞拍价格是显著偏大的。这一点反映出我们对于删失数据建模的重要性与合理性。

因为生存分析方法在不同领域具有普遍意义与普适应用<sup>[49]</sup>，本文介绍的基于深度神经网络的新型生存分析方法在医学、信息系统等其它领域也有巨大的应用价值。例如在医学领域，生存分析一般用于估计不同因素对于病人医疗效果或生存时间的影响；在线上信息系统中，生存分析还能用于预估用户下次访问的概率随时间的变化。

## 2.6 本章小结

在本章中，我们形式化定义了竞价形势预估的任务和建模面临的挑战，以及相关工作包括当时最先进模型的缺陷。为了解决这些问题与已有方法的缺陷，我们创新性地提出了基于深度生存分析方法的深度形式预估模型，对每一个交易样本做细粒度的建模与预估。我们的模型不仅摒弃了针对市场价格分布做预先假设的手段，还通过循环神经网络

络建模序列模式，更好地预估市场价格分布及其导出的胜率函数。实验分析印证了我们模型的有效性，实验结果可视化也从更深入的角度体现了模型的优异性能表现。

## 第三章 用户建模与效益预估

如前文图1-4所示，继上一章介绍实时竞价广告中的市场建模与成本预估之后，紧接着的研究课题是广告主端效益预估，其中涉及到用户建模、用户反馈预估与广告主效益优化。关于用户建模的内容将在第五章进行详细介绍。本章将形式化定义效益预估的核心问题，即用户反馈预估，并针对该问题介绍相关工作与现有工作存在的问题，并提出针对广告主期望收益进行优化的解决方案。

### 3.1 用户反馈预估

实时竞价广告中的效益预估，最核心的问题是用户反馈预估（User Response Prediction），即广告展示触达用户后，用户后续产生（正）反馈的行为概率，例如点击率（Click-through Rate, CTR）或购买转化率（Conversion Rate, CVR）。对于用户行为概率高的广告展示机会，一般广告主倾向于出相对较高的竞拍价格去获得本次广告展示；而对于用户反馈概率低的广告展示机会，广告主就倾向于出低价竞拍甚至不出价。这与经济学中的价值-价格关系一致。不失一般性，在本文中我们将以点击率预估为案例，其它类别的用户反馈预估，例如转化率预估等，均可以通过类似的方案去实现。

广告主的竞价决策优化过程是一个连续的过程。首先，用户的点击率预估一般被建模为二分类问题，即判断“用户点击与否”的概率，那么就可以用机器学习领域的分类模型<sup>[17, 50-52]</sup>进行建模与求解。这一阶段的主要目标是使得预测更加精准，例如尽可能减少预测值与真实值的差距。当广告主针对用户反馈进行预测之后，预测结果将作为输入提供给下一阶段，即竞价策略优化。竞价策略优化还需要考虑广告主的预算以及第二章所描述的市场竞争建模与市场价格预估。

事实上，这种序列化、分阶段的优化过程并非最优。根据贝叶斯决策理论<sup>[53]</sup>，用户反馈模型的优化更需要与后续竞价效益（bidding utility）相结合。举例来说，用户点击率预估的精度一般在  $[0, 1]$  之间有所不同，即不同概率区间的预测值分布可能导致最后效果有所差异。对于广告主来说，用户如果产生了点击，那么广告主的收益是正的；若用户没有产生点击，则广告主则会不会得到任何收益。同时，不同广告的用户点击价值又随着广告类型与广告内容的不同而有所不同。人们一般期望点击率预估模型能够更好地处理那些高点击率、高回报（即成本可能也很高）的广告展示机会，更精准地预测可能带来高回报的广告展示所对应的用户反馈。同时，由于实时竞价广告采用的是二价竞价机制，如前文章节2.1所述，出价最高的广告主赢下竞拍交易以后，他需要支付市

场价格，即市场第二高价；然后该广告主才能从用户点击或者转化等反馈行为中获得可能的收益回报。因此，市场价格以及市场竞争也对广告主最后收益有着重大的影响。如果广告主参与的竞拍交易竞争激烈，那么对于任意广告主来说，其赢得交易的置信度 (confidence) 就会降低。因此，相比竞争不太激烈的情形，用户点击率预估等效益预估模型在市场激烈竞争的情况下，更应该慎重考虑其预测行为，提高预测准确率。

综上所述，我们在这一章重新形式化定义了实时竞价广告场景中的用户反馈预估的问题，并以点击率预估为例进行阐述。和其它领域例如推荐系统中单纯将用户反馈预估建模为二分类问题不同的是，我们将用户反馈预估作为竞价优化的一部分，并且我们认为其预测输出将直接影响交易结果，也将最终影响广告主的广告投放结果；同时反过来，交易结果也会影响用户反馈预估的学习与优化。我们将直接优化广告主的广告活动收益并以此指导用户反馈预估模型的优化。以广告活动收益为优化目标，我们推导得出的梯度显示出用户反馈预估模型事实上是对市场竞争和市场价格分布有感知的。对于市场价格分布密度高的情况，其样本梯度的数值也高，这个优化目标也自然地针对市场竞争较为激烈的价格区间进行更多的学习。通过这种方式优化得到的模型将能在最终出价时非常接近真实市场价格，也就能给予广告主更加高效的竞拍策略，更加注重那些“高性价比”，即成本低，回报高的广告展示。值得注意的是，博弈论领域的相关文献已经证明<sup>[54]</sup>，在二价竞价机制下，广告主效益最优的出价策略是“真实出价” (truthful bidding) 策略，并与市场价格无关。然而一些经验研究<sup>[7]</sup> (empirical study) 表明，实时竞价广告中参与竞价的广告主往往是非理性的，他们的出价策略一般不是最优真实出价策略。所以在本文中我们和相关文献<sup>[26]</sup> 一样采取了一种随机的数学方法来建模出价策略。

在本文实验中，我们还构建了离线和在线真实竞价广告场景两种实验环境。我们的方法在离线环境下获得了 78.2% 的收益提升；而在线真实竞价广告场景下，相比传统的用户反馈预估模型，我们的方法获得了 25.5% 的广告收益提升。

本章剩余章节安排如下。第3.2节将详细介绍相关工作，包括用户反馈预估的现有方法及其缺陷、部分竞价策略优化的相关文献。第3.3节中我们将定义问题，详细介绍本文提出的基于广告主收益优化的用户反馈预估模型。在第3.4节，我们将详细阐述实验设置并分析实验结果。我们将在章节3.5总结本章内容并规划未来工作。

## 3.2 相关工作

### 3.2.1 用户反馈预估

用户反馈预估，例如估计用户点击率或转化率，已然成为实时展示广告研究领域的核心问题<sup>[18, 51, 55]</sup>。反馈预估建模了用户对广告主可能展示的广告内容的倾向性，并且被用于广告主后续预算分配与出价决策<sup>[50]</sup>的重要前序步骤。用户反馈预估一般被建模成一个二分类概率估计问题<sup>[56]</sup>，其训练目标则是最小化预测输出的似然概率与经验统计概率的偏差<sup>[1, 50, 52, 57]</sup>。从模型方法上来说，业界目前广泛采用了逻辑斯蒂回归<sup>[18]</sup>等线性模型、树模型<sup>[17]</sup>与因式分解机<sup>[52, 56]</sup>等非线性模型。另一些分类算法包括贝叶斯概率回归<sup>[1]</sup>、在线学习方法 FTRL (Follow the Regularized Leader)<sup>[58]</sup> 以及卷积神经网络方法<sup>[59]</sup>。最近，一些研究者指出<sup>[60, 61]</sup>，这些二分类模型的度量标准和评估方法难以合理地反映用户反馈预估模型在实时竞价广告场景下的真实性能，因为用户反馈预估在这个场景下是作为后续竞价策略使用的效益预估模块。所以我们在本文中重新建模和定义了基于广告主效益优化的模型框架，并在相关文献<sup>[60]</sup>的基础上更进一步优化了实时竞价广告场景下的用户反馈预估。

### 3.2.2 竞价策略

在本小节我们初步介绍实时竞价广告中广告主一般采用的竞价策略。在预估好点击率、转化率等用户反馈概率之后，广告主就能衡量一次广告展示所能带来的收益。有交易理论<sup>[62]</sup>证明了真实竞价是二价竞价机制下最优的竞价策略<sup>[18]</sup>。然而实际应用场景中，考虑到广告主预算以及市场上总交易数量的限制，真实竞价并非一定是最优的策略<sup>[9]</sup>。所以在真实交易场景中，线性竞价策略<sup>[63]</sup>是产业界广泛采用的一种出价策略。采用了这种出价策略的广告主将预估的点击率乘以一个常系数计算得到最终的出价，而这个常系数将根据广告活动的预算和广告投放的效果进行微调。另一篇最近发表的文章<sup>[64]</sup>是基于用户反馈概率提升方法的一种竞价策略。

然而，在所有这些相关工作中，市场价格分布或竞价形势从未被考虑在优化过程中。这种信息的缺失可能导致用户反馈预估偏离最后广告主的终极目标<sup>[60]</sup>，即广告投放的最终收益效果，所以目前所有基于二分类问题的用户反馈建模方案的模型都可能导致较差的竞价效果。在下文的实验中我们也证明了用户反馈预估和竞价策略进行协同优化，以及同时考虑市场价格分布的必要性和合理性。

### 3.3 直接优化广告收益的方法

在本章节，我们以点击率预估为例来详细定义用户反馈预估问题，并介绍我们提出的直接优化广告收益的方法。

#### 3.3.1 问题定义

在实时竞价广告中，每一位广告主直接或通过需求方代理平台对接了广告交易所。当用户发起网页浏览请求时，广告请求及其特征信息将被广告交易所转发给每一位参与竞价的广告主。广告主将竞价请求信息与其广告内容信息共同考虑并利用特征工程等进行预处理，得到竞价特征向量  $\mathbf{x}$ 。我们可以使用  $p_{\mathbf{x}}(\mathbf{x})$  来代表输入特征向量的分布。

点击率预估一般表示为函数  $p(y = 1|\mathbf{x}) \equiv f_{\theta}(\mathbf{x})$ ，该函数将特征向量  $\mathbf{x}$  映射成点击的概率，而点击行为变量  $y \in \{0, 1\}$  是一种二值变量，代表用户点击 (1) 或者没有点击 (0)。这个函数由  $\theta$  进行参数化。同时我们也用常量  $v$  来表示广告主得到用户点击后所能获得的正收益，该收益具体数值由广告主设定，一般可以通过广告商品价格乘以平均点击率获得。

接下来我们介绍点击率预估的上下文环境。一般现有工作将竞价策略通过一个函数  $b(f_{\theta}(\mathbf{x}))$  来表示。这个函数将预估所得的点击率（或者其它预估的用户正反馈概率）映射为具体出价<sup>[9, 18, 63]</sup>。一般地，这个映射过程服从序列依赖假设<sup>[9, 65]</sup>： $\mathbf{x} \rightarrow f_{\theta}(\mathbf{x}) \rightarrow b$ 。在本文中，我们也采用了这种形式化定义。同时，我们使用  $b(\cdot)$  来表示竞价函数，也用标量  $b$  来表示最终的竞拍价格。

当广告主发出竞拍价格  $b$  给广告交易所之后，广告交易所会进行交易处理<sup>[54]</sup>，决定出价最高的广告主赢得本次广告展示机会，如前文章节2.1所述。我们依旧使用  $z$  来表示市场价格，其概率密度函数为  $p_z(z)$ 。我们使用  $W(b)$  来表示胜率函数，即

$$W(b) = \int_0^b p_z(z) dz, \quad (3-1)$$

赢下交易的条件即为竞拍价格  $b$  大于市场价格  $z$ 。下文将阐述市场价格分布在点击率预估中的重要性。

若广告主赢下了竞拍交易，那么他将以市场价格  $z$  支付广告展示费用。所以我们用  $c(b)$  表示出价为  $b$  的时候，广告主的条件期望成本如下

$$c(b) = \frac{\int_0^b z p_z(z) dz}{\int_0^b p_z(z) dz}, \quad (3-2)$$

即广告主赢得交易的条件下支付的市场价格的期望<sup>[54]</sup>。

一旦我们确定了用户点击的回报  $v$ ，广告胜率  $W$  和期望成本  $c$ ，我们就能为数据集（所有参与竞拍并且赢下的广告交易样本集合）中的每一个交易样本二元组  $(\mathbf{x}, y)$  表示期望收益  $R_{\theta}(\mathbf{x}, y; b, v, c, W)$ 。

本章介绍的主要建模与优化任务即构建一种用户点击率（反馈）预估模型，通过机器学习的方法优化获得最优参数  $\theta^*$ ，使得期望收益最大化，即

$$\theta^* = \arg \max_{\theta} \int_{\mathbf{x}} R_{\theta}(\mathbf{x}, y; b, v, c, W) p_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}, \quad (3-3)$$

这里我们仅仅优化点击率预估模型，从而没有考虑广告主预算限制，因为在下文推导的公式中我们可以看到，点击率预估模型已经能够考虑到广告主成本  $c$ ，并控制其最大出价不超过用户点击的真实回报  $v$ ，即控制  $b < c$ 。考虑广告主的预算限制的模型即竞拍出价决策优化部分将在下一章重点介绍。

### 3.3.2 模型介绍

在本节，我们将以上文介绍的期望收益为优化目标，提出两种点击率预估模型的目标函数。在这两种目标函数的情况下，我们分别推导了优化参数的梯度，并讨论了两种梯度的性质及其与传统点击率预估优化方法的区别与优势。

#### 3.3.2.1 期望效益模型

针对公式 (3-3) 中的期望，我们首先提出了第一种期望效益（Expected Utility, EU）目标函数

$$R_{\theta}^{\text{EU}}(\mathbf{x}, y) = [vy - c(b(f_{\theta}(\mathbf{x})))] \cdot W(b(f_{\theta}(\mathbf{x}))), \quad (3-4)$$

为了简化符号，我们简化了  $R_{\theta}^{\text{EU}}(\mathbf{x}, y)$  中  $b, v, c, W$  的依赖关系。我们将训练集表示为  $D$ ，如前文所述，每一个样本是二元组  $(\mathbf{x}, y)$ ，其中  $\mathbf{x}$  是交易样本的特征向量， $y$  是用户点击指示符。那么所有交易的总体期望利润<sup>[60]</sup>可以通过将公式 (3-4) 代入公式 (3-1) 和 (3-2) 得到如下推导

$$\begin{aligned} & \sum_{(\mathbf{x}, y) \in D} R_{\theta}^{\text{EU}}(\mathbf{x}, y) \\ &= \sum_{(\mathbf{x}, y) \in D} \left[ vy - \frac{\int_0^{b(f_{\theta}(\mathbf{x}))} z \cdot p_z(z) dz}{\int_0^{b(f_{\theta}(\mathbf{x}))} p_z(z) dz} \right] \cdot \int_0^{b(f_{\theta}(\mathbf{x}))} p_z(z) dz \\ &= \sum_{(\mathbf{x}, y) \in D} \int_0^{b(f_{\theta}(\mathbf{x}))} (vy - z) \cdot p_z(z) dz. \end{aligned} \quad (3-5)$$

将公式 (3-5) 代入公式 (3-3) 并添加防止模型过拟合的正则项, 我们可以将优化目标转化为凸优化问题

$$\begin{aligned}\boldsymbol{\theta}^{\text{EU}} &= \arg \min_{\boldsymbol{\theta}} - \sum_{(\boldsymbol{x}, y) \in D} R_{\boldsymbol{\theta}}^{\text{EU}}(\boldsymbol{x}, y) + \frac{\lambda}{2} \|\boldsymbol{\theta}\|_2^2 \\ &= \arg \min_{\boldsymbol{\theta}} \sum_{\boldsymbol{x}} \int_0^{b(f_{\boldsymbol{\theta}}(\boldsymbol{x}))} (z - vy) \cdot p_z(z) dz + \frac{\lambda}{2} \boldsymbol{\theta}^T \boldsymbol{\theta},\end{aligned}\quad (3-6)$$

其中参数  $\boldsymbol{\theta}$  的最优结果将通过梯度下降的方法得到。接下来我们推导参数  $\boldsymbol{\theta}$  所对应的优化目标  $R_{\boldsymbol{\theta}}^{\text{EU}}(\boldsymbol{x}, y)$  的梯度如下

$$\begin{aligned}\frac{\partial R_{\boldsymbol{\theta}}^{\text{EU}}(\boldsymbol{x}, y)}{\partial \boldsymbol{\theta}} &= \overbrace{(b(f_{\boldsymbol{\theta}}(\boldsymbol{x})) - vy)}^{\text{竞价误差}} \cdot \overbrace{p_z(b(f_{\boldsymbol{\theta}}(\boldsymbol{x})))}^{\text{市场敏感性}} \\ &\quad + \frac{\partial b(f_{\boldsymbol{\theta}}(\boldsymbol{x}))}{\partial f_{\boldsymbol{\theta}}(\boldsymbol{x})} \frac{\partial f_{\boldsymbol{\theta}}(\boldsymbol{x})}{\partial \boldsymbol{\theta}} + \lambda \boldsymbol{\theta},\end{aligned}\quad (3-7)$$

在实际的模型学习过程中, 我们通过以上的推导结果, 利用随机梯度下降进行参数优化:  $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} - \eta \frac{\partial R_{\boldsymbol{\theta}}^{\text{EU}}(\boldsymbol{x}, y)}{\partial \boldsymbol{\theta}}$ 。

**模型讨论.** 公式 (3-7) 为点击率预估提供了一种新的优化思路。EU 模型通过将广告投放效益和竞价决策的成本 (竞价误差, **Bid Error**) 进行协同考虑, 同时也将市场价格分布 (市场敏感性, **Market Sensitivity**) 归纳在梯度函数中。梯度中这两个部分扮演着权重重新分配 (**re-weighting**) 的功能, 使得点击率预估模型的梯度能够感知到出价策略及市场竞争的影响。我们在图3-1中展示了梯度中这两个部分的影响。左图展示了竞价误差 **Bid Error** 部分在不同竞拍价格与不同用户反馈时的数值变化。通过竞拍价格  $b(f_{\boldsymbol{\theta}}(\boldsymbol{x}))$  与实际收益  $vy$  之间的差异, 点击率预估模型的梯度会根据训练样本的实际收益进行梯度微调, 也就是说, 一个训练完善的模型 (参数) 会针对正反馈 (用户产生点击的) 样本生成非常接近广告主真实收益  $v$  的出价; 而针对无反馈 (用户未产生点击的) 样本生成接近于零值的出价。右图绘制了市场敏感性 **Market Sensitivity** 模块 (左 Y 轴) 对梯度的影响, 以及竞价误差和市场敏感性两部分的协同影响 (右 Y 轴)。我们观察到市场敏感性通过和预估的市场价格分布 (上文第二章介绍) 进行适应性计算来纠正用户竞价差异, 这样就使得用户点击率模块在学习时能够对市场价格分布相接近的出价区间给予更多的关注, 即市场价格概率密度越高的地方, 梯度修正效果越强。从直观上来说, 这种修正方式是非常合理的, 因为广告主在训练模型时更应该关注与市场价格接近且市场价格分布更加密集的价格区间, 这样能够针对市场竞争激烈的价格区间进行预测值的精细微调, 达到提升竞拍效益的效果。

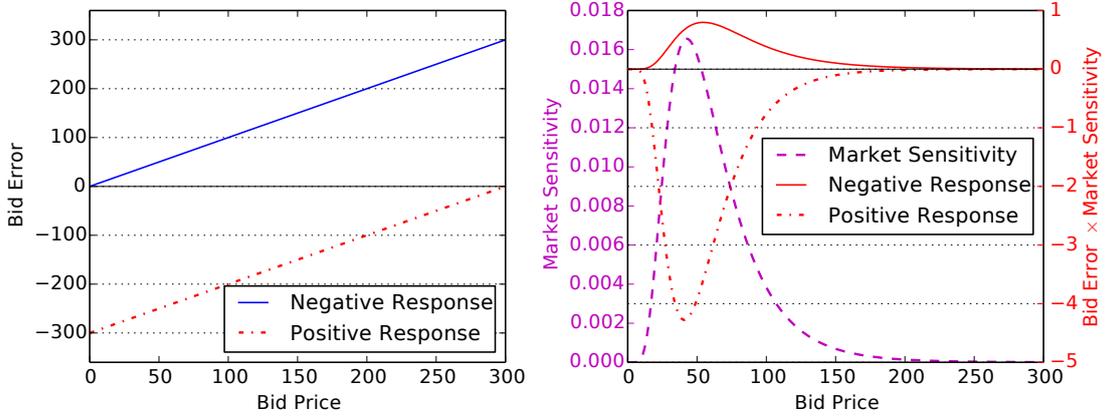


图 3-1 期望效益模型的梯度分析；用户点击收益为  $v = 300$ 。

Fig 3-1 The impact illustration from the bid and market price of Expected Utility (EU); click value  $v = 300$ .

### 3.3.2.2 风险回报模型

除了上一节提出的期望效益 EU 模型，我们还提出了一种风险回报 (Risk Return, RR) 模型，它通过调整出价决策中风险与回报的平衡来实现效益优化，即

$$R_{\theta}^{RR}(\mathbf{x}, y) = \left( \underbrace{\frac{vy}{z}}_{\text{回报}} - \underbrace{\frac{v(1-y)}{v-z}}_{\text{风险}} \right) \cdot W(b(f_{\theta}(\mathbf{x}))). \quad (3-8)$$

在风险回报模型中，当用户反馈为正反馈（点击），即  $y = 1$  时，公式 (3-8) 中的效益是  $\frac{v}{z}$ ，代表赢得竞拍交易后回报与广告主成本的比值；当用户没有正反馈，即  $y = 0$  时，效益则变成了决策风险  $\frac{-v}{v-z}$ ，也即赢得以后没有回报的损失  $(-v)$  与赢得以后的回报  $(v-z)$  之间的比值。注意到广告主的回报总是高于市场价格  $z$  的，即  $v \geq b > z$ 。所以当广告主的回报与市场价格比较接近的时候（分母部分  $(v-z)$  较小），对于风险与负回报的惩罚是相对较高的。

因此，风险回报模型的目标函数则可以写成

$$\begin{aligned} \theta^{RR} &= \arg \min_{\theta} - \sum_{(\mathbf{x}, y) \in D} R_{\theta}^{RR}(\mathbf{x}, y) + \frac{\lambda}{2} \|\theta\|_2^2 \\ &= \arg \min_{\theta} - \sum_{(\mathbf{x}, y) \in D} \int_0^{b(f_{\theta}(\mathbf{x}))} \left( \frac{vy}{z} - \frac{v(1-y)}{v-z} \right) \cdot p_z(z) dz \\ &\quad + \frac{\lambda}{2} \theta^T \theta, \end{aligned} \quad (3-9)$$

这也能够推导得到风险回报模型参数的梯度如下

$$\frac{\partial R_{\theta}^{\text{RR}}(\mathbf{x}, y)}{\partial \theta} = \left( \overbrace{-\frac{vy}{b(f_{\theta}(\mathbf{x}))} + \frac{v(1-y)}{v-b(f_{\theta}(\mathbf{x}))}}^{\text{竞价误差}} \right) \cdot \overbrace{p_z(b(f_{\theta}(\mathbf{x})))}_{\text{市场敏感性}} \cdot \frac{\partial b(f_{\theta}(\mathbf{x}))}{\partial f_{\theta}(\mathbf{x})} \frac{\partial f_{\theta}(\mathbf{x})}{\partial \theta} + \lambda \theta. \quad (3-10)$$

**讨论.** 为了更好地理解梯度的含义与性质，我们在图3-2中绘制了风险回报模型中竞价误差与市场敏感性两部分对梯度的影响。风险回归模型和之前介绍的期望效益模型不同，前者的竞价误差会在用户反馈为正反馈 ( $y = 1$ ) 时转化为“回报”；并会在用户没有正反馈 ( $y = 0$ ) 时转化为“风险”。当  $y = 0$  并且竞拍价格较高，或者当  $y = 1$  但是竞拍价格较低时，梯度中竞拍误差部分的重要性就凸显了出来，它会防止用户反馈预估模型出现这两种情况。

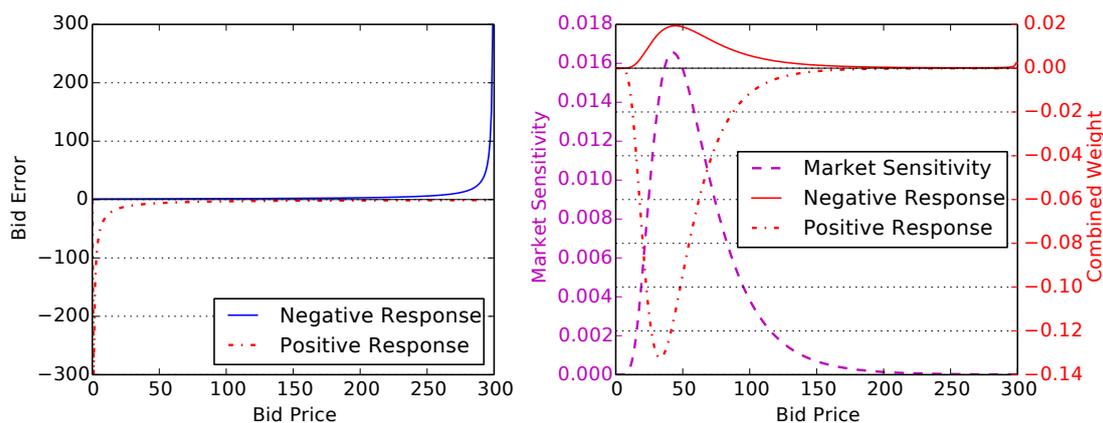


图 3-2 风险回报模型的梯度分析；用户点击收益为  $v = 300$ 。

Fig 3-2 The impact illustration from the bid and market price of Risk Return (RR); click value  $v = 300$ .

正如公式 (3-7) 和 (3-10) 所示，市场价格分布，即上一章讨论的竞价形势，在用户反馈预估优化过程中扮演了重要的角色：当用户反馈预估模型和竞价函数确定的时候，用户反馈模型的梯度受市场价格分布  $p_z(z)$  的值的影响。

不同的竞价形势预估模型可以用在这里计算市场价格分布  $p_z(z)$ ，例如参数化模型<sup>[24, 60]</sup>。在本章节，为了灵活建模不同的市场价格分布并分析我们新提出的用户反馈预估模型的特点，我们采用了一种非参数化的统计方法建模市场价格分布，即直接统计每一个广告活动数据集中的不同市场价格频率。我们将在第四章介绍直接用于优化广告投放收益的竞价形势预估模型。

### 3.3.3 模型分析

上述公式 (3-6) 和 (3-9) 中的两种点击率预估模型的学习过程需要依赖竞价函数  $b(f_{\theta}(\mathbf{x}))$ 、市场价格分布函数  $p_z(z)$  以及自身的点击率预估函数  $f_{\theta}(\mathbf{x})$ 。在本节，我们将讨论以上函数在特定的几种形式下，点击率预估模型优化方法的最终结果以及与传统点击率预估模型方法的相似性与不同之处。

不失一般性，我们为点击率预估函数设定为逻辑斯蒂回归函数的形式：

$$f_{\theta}(\mathbf{x}) \equiv \sigma(\boldsymbol{\theta}^T \mathbf{x}) = \frac{1}{1 + e^{-\boldsymbol{\theta}^T \mathbf{x}}}, \quad (3-11)$$

该函数形式是目前产业界应用最多的点击率预估模型。其他模型包括深度神经网络等模型亦可用于  $f$  函数的拟合，但是不同模型的实现不是本章讨论重点，故此略过。

之后，我们就能计算得到函数  $f$  关于模型参数  $\boldsymbol{\theta}$  的梯度  $\frac{\partial f_{\theta}(\mathbf{x})}{\partial \boldsymbol{\theta}} = \sigma(\boldsymbol{\theta}^T \mathbf{x})(1 - \sigma(\boldsymbol{\theta}^T \mathbf{x}))\mathbf{x}$ 。

对于竞价策略函数，我们采用了广泛采用的线性竞价函数，即出价与预估的点击率成正比关系<sup>[63]</sup>，比例系数为  $\gamma$ ：

$$b(f_{\theta}(\mathbf{x})) \equiv \gamma \cdot v \cdot f_{\theta}(\mathbf{x}). \quad (3-12)$$

将公式 (3-11) 和 (3-12) 分别代入公式 (3-7) 和 (3-10) 我们可以得到最终的期望效益模型的优化梯度：

$$\begin{aligned} \frac{\partial R_{\theta}^{\text{EU}}(\mathbf{x}, y)}{\partial \boldsymbol{\theta}} &= \gamma v^2 (\gamma \sigma(\boldsymbol{\theta}^T \mathbf{x}) - y) \cdot p_z(b(f_{\theta}(\mathbf{x}))) \cdot \\ &\quad \sigma(\boldsymbol{\theta}^T \mathbf{x})(1 - \sigma(\boldsymbol{\theta}^T \mathbf{x}))\mathbf{x} + \lambda \boldsymbol{\theta}, \end{aligned} \quad (3-13)$$

和风险回报模型的优化梯度：

$$\begin{aligned} \frac{\partial R_{\theta}^{\text{RR}}(\mathbf{x}, y)}{\partial \boldsymbol{\theta}} &= \gamma v \left( -\frac{y}{\gamma \sigma(\boldsymbol{\theta}^T \mathbf{x})} + \frac{1 - y}{1 - \gamma \sigma(\boldsymbol{\theta}^T \mathbf{x})} \right) \cdot \\ &\quad p_z(b(f_{\theta}(\mathbf{x}))) \cdot \sigma(\boldsymbol{\theta}^T \mathbf{x})(1 - \sigma(\boldsymbol{\theta}^T \mathbf{x}))\mathbf{x} + \lambda \boldsymbol{\theta}, \end{aligned} \quad (3-14)$$

其中， $\gamma$  在出价函数中承担着一种校准 (calibration) 的作用，用于真实出价部分 ( $v \cdot f_{\theta}(\mathbf{x})$ ) 的微调<sup>[9, 63]</sup>。

本文提出的用户反馈预估模型直接优化广告主收益和广告投放效益，所以接下来我们将我们的模型及其梯度与传统基于二分类方法的模型进行对比。逻辑斯蒂回归模型一般可以用平方误差 (Squared Error) 作为损失函数进行监督训练，其损失函数和对应的

梯度为:

$$\begin{aligned}\mathcal{L}_{\theta}^{\text{SE}}(\mathbf{x}, y) &= \frac{1}{2}(y - \sigma(\boldsymbol{\theta}^T \mathbf{x}))^2, \\ \frac{\partial \mathcal{L}_{\theta}^{\text{SE}}(\mathbf{x}, y)}{\partial \boldsymbol{\theta}} &= (\sigma(\boldsymbol{\theta}^T \mathbf{x}) - y)\sigma(\boldsymbol{\theta}^T \mathbf{x})(1 - \sigma(\boldsymbol{\theta}^T \mathbf{x}))\mathbf{x}.\end{aligned}\quad (3-15)$$

在分类任务中,更为通用的逻辑斯蒂回归模型的目标函数是交叉熵损失函数,其损失函数和对应的梯度为

$$\begin{aligned}\mathcal{L}_{\theta}^{\text{CE}}(\mathbf{x}, y) &= -y \log \sigma(\boldsymbol{\theta}^T \mathbf{x}) - (1 - y) \log(1 - \sigma(\boldsymbol{\theta}^T \mathbf{x})), \\ \frac{\partial \mathcal{L}_{\theta}^{\text{CE}}(\mathbf{x}, y)}{\partial \boldsymbol{\theta}} &= (\sigma(\boldsymbol{\theta}^T \mathbf{x}) - y)\mathbf{x}.\end{aligned}\quad (3-16)$$

我们可以发现,在本文提出的优化方法,即公式(3-13)和(3-14),将公式(3-15)和(3-16)中的原始梯度进行了拓展:

- 将用户反馈误差用竞价差异代替;
- 在梯度更新中考虑了不同竞拍价格所处的市场竞争情况,即将竞价形势考虑在内。

接下来我们将分别进行讨论。

**用户反馈误差对比竞价误差的讨论.**在实时竞价广告场景下,直接优化(最小化)竞价误差损失相比最小化用户反馈误差更加有效,因为广告主可能根本无法获得一个点击的真实价值,即当广告主未赢下广告竞拍时,根本无从谈起用户反馈  $y$  的具体真值。我们的解决方案使用了线性竞价函数,并能够很好地弥补竞拍出价与真实广告市场价格的差异并相应地通过反馈预估模型进行出价微调。为了达到类似的效果,很多广告主确切知晓用户正反馈的真实值  $v$  并采用了真实竞价策略函数<sup>[18, 63]</sup>:

$$b(f_{\theta}(\mathbf{x})) = v \cdot f_{\theta}(\mathbf{x}). \quad (3-17)$$

如果广告主使用真实竞价函数并设置点击率预估函数为逻辑斯蒂回归函数,那么期望效应模型与风险回报模型的梯度(如公式(3-7)和(3-10)所示)将分别简化为

$$\begin{aligned}\frac{\partial R_{\theta}^{\text{EU}}(\mathbf{x}, y)}{\partial \boldsymbol{\theta}} &= v^2(\sigma(\boldsymbol{\theta}^T \mathbf{x}) - y) \cdot p_z(b(f_{\theta}(\mathbf{x}))) \\ &\quad \cdot \sigma(\boldsymbol{\theta}^T \mathbf{x})(1 - \sigma(\boldsymbol{\theta}^T \mathbf{x}))\mathbf{x} + \lambda \boldsymbol{\theta},\end{aligned}\quad (3-18)$$

$$\frac{\partial R_{\theta}^{\text{RR}}(\mathbf{x}, y)}{\partial \boldsymbol{\theta}} = v(\sigma(\boldsymbol{\theta}^T \mathbf{x}) - y)p_z(b(f_{\theta}(\mathbf{x})))\mathbf{x} + \lambda \boldsymbol{\theta}, \quad (3-19)$$

其中期望效益模型的误差部分和平方误差损失的模型梯度(公式(3-15))中的误差部分一致,而风险回报的误差部分与交叉熵损失的模型梯度(如公式(3-16))中的误差部分一致。但是他们又分别受市场敏感性  $p_z(b(f_{\theta}(\mathbf{x})))$  的权重所影响。

表 3-1 两种用户反馈预估模型与传统方法的梯度对比, LR: 逻辑斯蒂回归, TB: 真实出价策略, LB: 线性出价策略, UM: 均一市场价格分布。LR 和 LR+TB+UM 是等价的 (当假设市场价格分布是均一分布时, LR+TB 退化为 LR)

Table 3-1 The comparison of the model gradients (without regularization). LR: logistic regression, TB: truthful bidding, LB: linear bidding, UM: uniform market price distribution. LR and LR+TB+UM are equivalent (LR+TB reduces to the baseline LR when assuming the uniform market price distribution).

模型设置	EU (SE) 梯度	RR (CE) 梯度
LR (基线模型)	$\frac{\partial C_{\theta}^{SE}(\mathbf{x}, y)}{\partial \theta} = (\sigma(\theta^T \mathbf{x}) - y) \cdot \sigma(\theta^T \mathbf{x})(1 - \sigma(\theta^T \mathbf{x}))\mathbf{x}$	$\frac{\partial C_{\theta}^{RR}(\mathbf{x}, y)}{\partial \theta} = (\sigma(\theta^T \mathbf{x}) - y)\mathbf{x}$
LR+TB	$\frac{\partial R_{\theta}^{EU}(\mathbf{x}, y)}{\partial \theta} = v^2(\sigma(\theta^T \mathbf{x}) - y) \cdot p_z(b(f_{\theta}(\mathbf{x}))) \cdot \sigma(\theta^T \mathbf{x})(1 - \sigma(\theta^T \mathbf{x}))\mathbf{x}$	$\frac{\partial R_{\theta}^{RR}(\mathbf{x}, y)}{\partial \theta} = v(\sigma(\theta^T \mathbf{x}) - y) \cdot p_z(b(f_{\theta}(\mathbf{x}))) \cdot \mathbf{x}$
LR+TB+UM	$\frac{\partial R_{\theta}^{EU}(\mathbf{x}, y)}{\partial \theta} = v^2 l(\sigma(\theta^T \mathbf{x}) - y) \cdot \sigma(\theta^T \mathbf{x})(1 - \sigma(\theta^T \mathbf{x}))\mathbf{x}$	$\frac{\partial R_{\theta}^{RR}(\mathbf{x}, y)}{\partial \theta} = vl(\sigma(\theta^T \mathbf{x}) - y)\mathbf{x}$
LR+LB	$\frac{\partial R_{\theta}^{EU}(\mathbf{x}, y)}{\partial \theta} = \gamma v^2(\gamma \sigma(\theta^T \mathbf{x}) - y) \cdot p_z(b(f_{\theta}(\mathbf{x}))) \cdot \sigma(\theta^T \mathbf{x})(1 - \sigma(\theta^T \mathbf{x}))\mathbf{x}$	$\frac{\partial R_{\theta}^{RR}(\mathbf{x}, y)}{\partial \theta} = \gamma v \left( -\frac{y}{\gamma \sigma(\theta^T \mathbf{x})} + \frac{1-y}{1-\gamma \sigma(\theta^T \mathbf{x})} \right) \cdot p_z(b(f_{\theta}(\mathbf{x}))) \cdot \sigma(\theta^T \mathbf{x})(1 - \sigma(\theta^T \mathbf{x}))\mathbf{x}$

关于市场价格分布的讨论. 需要注意的是, 我们在模型中采用了非参数化的市场价格分布建模方法, 直接从广告历史数据中进行统计。在这里我们考虑一种极其简化的特殊情况, 即假设市场价格服从均一分布:

$$p_z(z) = l, \tag{3-20}$$

故公式 (3-18) 能够简化成

$$\frac{\partial R_{\theta}^{EU}(\mathbf{x}, y)}{\partial \theta} = v^2 l(\sigma(\theta^T \mathbf{x}) - y) \cdot \sigma(\theta^T \mathbf{x})(1 - \sigma(\theta^T \mathbf{x}))\mathbf{x} + \lambda \theta \tag{3-21}$$

这也就和传统的平方误差逻辑斯蒂回归模型的梯度完全一致, 如公式 (3-15) 所示。而公式 (3-19) 则简化成

$$\frac{\partial R_{\theta}^{RR}(\mathbf{x}, y)}{\partial \theta} = vl(\sigma(\theta^T \mathbf{x}) - y)\mathbf{x} + \lambda \theta, \tag{3-22}$$

和传统的交叉熵逻辑斯蒂回归模型的梯度完全一致, 如公式 (3-16) 所示。

综上所述, 我们能够发现在假设 (1) 出价策略为真实出价策略; (2) 市场价格服从均一分布时, 我们提出的直接优化广告主效益与利润的用户反馈预估模型同传统的二分类逻辑斯蒂回归模型 (平方误差与交叉熵损失函数) 有相同的梯度形式。表3-1总结和展示了在不同设定下, 我们的期望效益模型与风险回报模型的梯度形式。需要注意的是, 以上分析基于一些简单的假设, 而在实际模型应用过程中, 我们采用了更加复杂、合理, 更加符合真实情况的出价函数和市场价格分布, 在实验中我们的两种模型相比传统模型取得了巨大的效果提升。

### 3.4 实验分析

在本章, 我们将详细介绍实验设置、对比模型、评估方法并分析具体的实验结果。

### 3.4.1 数据集与实验流程

**数据集.** 同第二章一致, 我们使用了 **iPinYou** 和 **YOYI** 两个数据集。详细的数据集介绍可以参考本文章节2.5.1。另外, 在实时竞价广告数据集中, 正例 (用户点击过的) 样本相比负例 (用户未产生点击的) 样本稀疏很多, 所以我们和相关文献<sup>[17]</sup> 一样采用了负采样的数据处理方案, 并为每个模型输出使用了相同的校准方法, 使得用户点击行为的样本空间正负比例与真实数据来源保持一致。

另外, 我们还基于 **YOYI** 公司真实的实时竞价广告平台构建了在线实验环境, 用以在真实场景中实验测试我们方法的有效性。

**实验流程.** 我们将原始广告展示数据作为离线竞价请求集合。经过特征工程等预处理以后, 数据集包含一系列竞价记录三元组  $(\mathbf{x}, z, y)$ , 即竞价请求特征向量  $\mathbf{x}$ , 交易市场价格  $z$  和用户反馈标记  $y \in \{0, 1\}$ 。我们和相关文献<sup>[9]</sup> 采用了一样的实验流程: 每一个竞价历史的特征向量根据历史上产生的时间顺序被当做离线竞价请求被发送给离线竞价引擎, 离线竞价引擎会复现当时的竞价过程, 即根据收到的竞价请求计算收益与成本并根据竞价策略进行出价。若模型给出的竞拍价格大于数据样本中的市场价格, 则记为该模型赢得本次离线广告展示, 否则认为没有赢得该广告展示机会。一方面, 我们给竞价引擎装入不同的点击率预估模型来预测用户点击概率, 并进行性能对比; 另一方面, 我们给所有点击率预估模型配置相同的出价策略函数, 保持实验对比的公平性。在离线竞价过程全部结束后, (不同点击率预估模型对应的) 离线竞价引擎获得的所有用户点击与累积成本被汇总起来, 以备后续实验指标的计算。这个离线评估流程已被许多相关工作<sup>[9, 26]</sup> 采用, 用于评估用户反馈预估模型与竞价策略优化的各类解决方案的性能, 也被产业界广泛采用<sup>[45]</sup>。

### 3.4.2 评估指标与对比模型

**评估指标.** 因为我们的优化目标是提高广告活动的最终效益, 并尽量减少广告投放的成本, 所以在性能评估时, 我们主要考虑竞价过程结束后, 离线竞拍所获得的**利润 (profit)** 与**投资回报率 (Risk of Investment, ROI)**。当离线竞拍引擎在实时竞价交易中赢得一次广告展示机会时, 广告主会支付对应交易的市场价格, 支出会被累加到总**成本 (cost)** 中。当该交易对应的用户反馈 (点击) 为正反馈时, 我们会将用户点击收益常量  $v$  加总到总**收益 (return)** 中。这个广告收益  $v$  的具体值一般由广告主确定, 我们将训练数据集中统计得到的有效点击费用 (effective Cost per Click, eCPC), 即训练集中所有样本市场价格的和除以总点击数目的比值, 作为测试过程中的点击收益常量, 这个数值对于所有模型都是相等的。**利润**被当做是测试样本集合中的总的毛利 ( $\sum \text{return} - \sum \text{cost}$ ); 投资回

报率则是反应竞价策略的效益的另一个重要指标，**ROI** 的计算方法是  $\text{profit}/\text{cost}$ 。

为了对比用户反馈预估的分类效果，即正确预测用户反馈概率（点击率）的能力，我们和其他用户反馈预估相关工作一样采用了 **AUC** (Area Under ROC Curve) 和 **RMSE** (Root Mean Squared Error) 来衡量分类效果。我们还使用了基于测试数据集计算得出的 **eCPC** (越低越好)、每千次展示成本 (cost per thousand impressions, **CPM**)、平均点击率 (**CTR**) 和平均胜率 (**Winning Rate**) 来对比不同用户反馈预估模型支持的竞价策略的性能优劣。

**对比模型.** 在实验中，我们对分析四个不同的模型。

**CE** 是被广泛采用的逻辑斯蒂回归模型<sup>[17, 51]</sup>，用于实时竞价广告、推荐系统等任务中。这个模型采用交叉熵作为其优化目标函数，其参数的梯度如公式 (3-16) 所示。

**SE** 是采用平方误差作为损失函数的逻辑斯蒂回归方法，其参数的梯度如公式 (3-15) 所示。

**EU** 是我们的期望效益模型，其参数梯度如公式 (3-13) 所示。

**RR** 是我们的风险回报模型，其参数梯度如公式 (3-14) 所示。

需要注意的是，选择这些对比模型的依据是其优化的损失函数有所区别，而我们限制了模型本身采用一致的方案，即点击率预估模型  $f_{\theta}(\mathbf{x})$  使用逻辑斯蒂回归模型，并且不失一般性。其它模型例如深度学习方法亦可通过替换  $f$  函数的实现进行对比，这里我们不做展开讨论。

### 3.4.3 实验结果与对比分析

**预估准确性实验.** 本节将对比不同点击率预估模型的预估准确性，通过 **AUC** 和 **RMSE** 两个指标体现。因为点击率预估模型的根本任务是估准用户点击概率，即正反馈概率；而我们提出的两种新型点击率预估模型的优化目标是广告主的收益。为了对比我们的模型和传统模型在点击率预估方面的分类准确率方面的性能，我们在本节进行详细的实验对比与分析。表3-2展现了所有对比模型在 **AUC** 和 **RMSE** 两个指标上的性能结果。

首先，我们从表格中可以看到采用交叉熵损失函数的 **CE** 模型比另一个基线模型性能更好，后者即采用平方误差损失函数的 **SE** 模型。这个现象也解释了二分类问题中研究者更倾向于采用交叉熵损失函数，而基于平方误差损失函数的模型更多应用于输出值为连续值的回归问题<sup>[66]</sup>。其次，我们在文中提出的两个模型，**EU** 模型和 **RR** 模型均获得了与基线 **CE** 模型相似甚至更好的 **AUC** 指标评测结果，同时在 **RMSE** 中保持了差不多的性能水平。回顾章节3.3.3的分析，**EU** 模型相比 **SE** 模型的重要优势在于模型在优化时，其梯度还考虑了市场价格分布的信息。在表格中我们可以看到，**EU** 模型不仅弥

表 3-2 用户反馈预估模型的回归性能对比, AUC 越高性能越好; RMSE 越低性能越好

Table 3-2 Regression performances, AUC: the higher, the better. RMSE: the smaller, the better

iPinYou	AUC				RMSE ( $\times 10^{-2}$ )			
	SE	CE	EU	RR	SE	CE	EU	RR
1458	.948	.987	.987	.977	3.01	1.94	2.42	2.32
2259	.542	.692	.674	.691	2.01	1.77	1.76	1.79
2261	.490	.569	.622	.619	1.84	1.68	1.71	1.68
2821	.511	.620	.608	.639	2.56	2.43	2.39	2.46
2997	.543	.610	.606	.608	5.98	5.82	5.84	5.82
3358	.863	.974	.970	.980	3.07	2.47	3.32	2.67
3386	.593	.768	.761	.778	2.95	2.84	3.32	2.85
3427	.634	.976	.976	.960	2.78	2.20	2.61	2.34
3476	.575	.957	.954	.950	2.50	2.32	2.39	2.33
平均	.633	.794	.795	<b>.800</b>	2.97	<b>2.61</b>	2.86	2.69
YOYI	.882	.891	<b>.912</b>	<b>.912</b>	11.9	11.7	11.8	<b>11.6</b>

补了 SE 模型在分类准确率上 (相比另一基线模型 CE) 的劣势, 还在 2261 和 YOYI 两个子数据集上获得了更好的结果。同时可以注意到, EU 模型和相对较优的 CE 模型拥有相似的分类效果, 甚至在某些子数据集上更好。最后, 我们还观察到 RR 模型在大部分广告活动子数据集上表现稳定, 同时取得了相对最好的分类效果, 例如 2821、3358、3386 和 YOYI 数据集, 这也表明了采用交叉熵的逻辑斯蒂回归模型结合市场价格分布的感知是最有效的分类组合, 虽然二者结合的目的是为了优化广告投放的效益。

**离线广告投放利润分析.** 在本节我们分析不同点击率预估模型驱动的竞价策略在离线竞拍模拟过程中的收益表现。图3-3展示了模型的学习曲线, 横坐标是不同模型在整个训练集合上的完整迭代次数 (round), 纵坐标则是广告主的收益。

首先我们对比了两个基线模型 (SE 与 CE) 的利润与 ROI 效果, 结果列在表3-3中。从表格结果中我们可以发现, 两个基线模型在几乎所有广告活动数据集中都取得了正 ROI。同时在大部分广告活动中, 除了 2821, CE 模型都比 SE 模型取得了更好的利润收益与 ROI 效果, 这个和上一节中我们的发现一致, 即 CE 模型比 SE 模型拥有更优秀的分类性能。

我们接下来选取表现相对较好的基线模型 CE 作为基准, 对比我们提出的两种模型, 表3-4列出了详细的评估结果。从表中我们可以发现: (1) 相比 CE 模型, 我们提出的 EU 和 RR 模型几乎都获得了一致的利润的提升, 只有在 3386 中 EU 获得了相对较少的利润。不过综合来看, 这两个新模型分别提升了 71.2% 和 78.2% 的利润增益。(2)

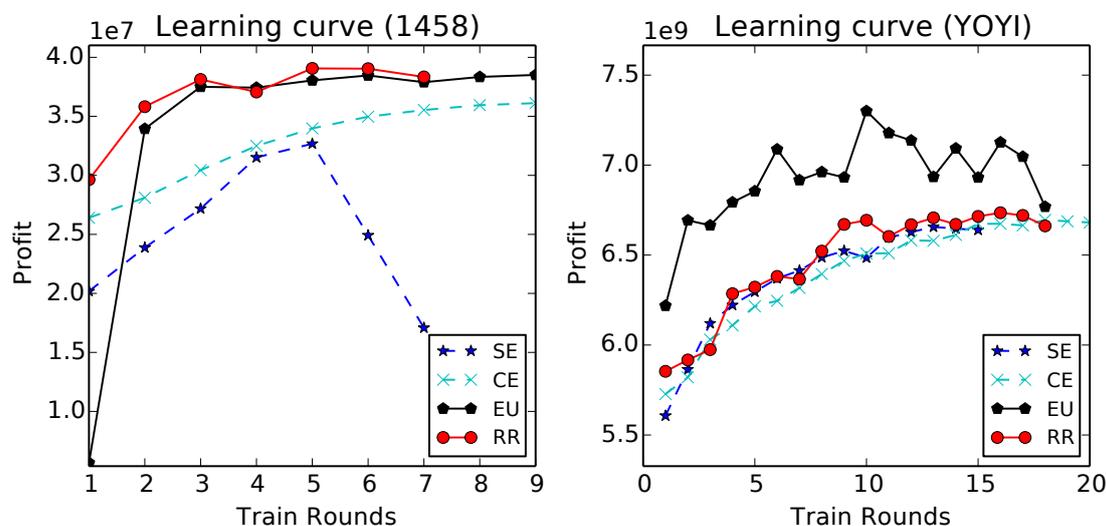


图 3-3 用户反馈预估模型的学习曲线

Fig 3-3 Learning curve on iPinYou (left) and YOYI (right)

在 ROI 指标结果中，EU 和 RR 模型相比较基线模型获得了更大的提升，分别达到 202% 和 217% 的增益。这也显示出我们提出的模型能够显著降低消耗，提升毛利率和投资回报，他们能够更加关注高回报（高用户正反馈概率）低成本（低市场价格）的那些广告展示。(3) 相对来说，RR 是表现最优的模型，相比 EU 提升了 7.0% 的利润和 15% 的 ROI。

最后，表3-5列出了离线实时竞价实验中四个对比模型在其他评估指标下的结果。从表中我们能够得出以下分析结论：(1) 点击率 CTR 反映了赢得的广告展示机会的总体质量，若点击率越高，则说明模型分类准确率越高，能够精准地定位那些高点击概率的广告展示机会。我们的模型，包括 EU 和 RR，均获得了比基线模型更高的点击率性能。(2) CPM 体现了赢下广告展示的总体成本，我们的模型和基线模型有着相似的成本水平。(3) 结合点击率和 CPM，我们能够计算得出 eCPC 指标，这个指标也是和利润与 ROI 最直接相关的度量标准，即 eCPC 越低，ROI 越高，代表竞价策略所获得的广告展示机会性价比越高，广告主的效益越高。我们提出的两个模型 EU 和 RR 在绝大部分广告活动数据集中均获得了较低（较好）的 eCPC 值，这也解释了为什么我们的模型能够获得更多的利润与更高的 ROI。(4) EU 和 RR 模型的胜率较低，这也反映出我们的模型能够避免将预算过多浪费在低回报（低用户反馈概率）的广告展示交易中。

**出价决策行为分析.** 在本节，我们深入分析了不同点击率预估模型驱动的出价策略的竞价行为，阐述本章实验结果背后的原因。正如之前的章节内容所述，我们提出的点击率预估模型的核心优势在于其学习过程中对于市场价格分布与竞价误差的感知。为了更直

表 3-3 广告活动利润效果对比 (基线模型)

Table 3-3 Direct campaign profit of baselines

iPinYou	利润 ( $\times 10^7$ )		ROI	
	SE	CE	SE	CE
1458	3.2	3.6	4.2	6.6
2259	-0.32	0.40	-0.080	0.18
2261	0.29	0.63	0.26	0.40
2821	0.11	0.08	0.21	0.023
2997	0.11	0.14	0.42	0.71
3358	1.76	2.4	5.4	5.2
3386	0.51	1.6	0.16	1.2
3427	0.33	2.9	0.11	3.4
3476	0.65	3.1	0.36	3.5
平均	0.74	<b>1.7</b>	1.2	<b>2.3</b>
YOYI	665.6	<b>669.5</b>	1.8	<b>1.9</b>

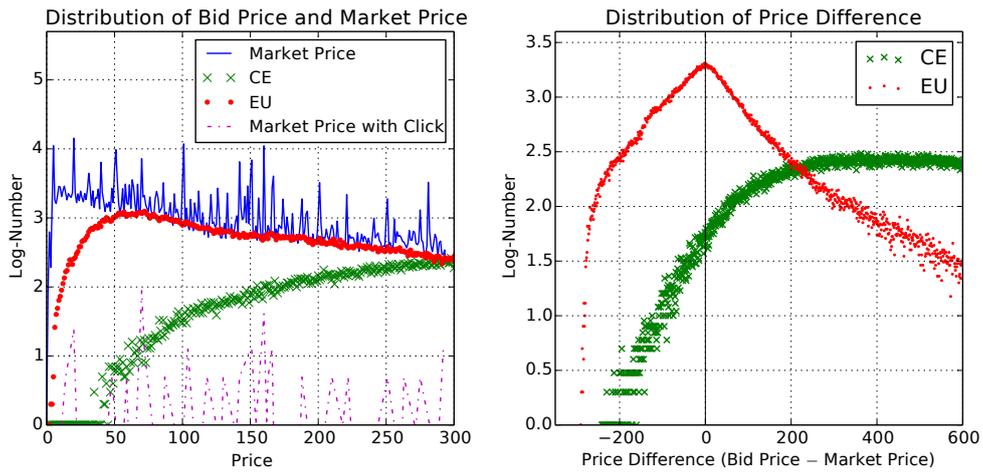


图 3-4 模型输出的竞拍价格分布示意图

Fig 3-4 Analysis of bid price and market price distribution (iPinYou campaign 2259)

观地理解市场价格分布在梯度中的作用，我们在图3-4中的左图分别绘制了我们的 EU 模型和基线模型 CE 驱动的出价策略的竞拍价格分布，并且将他们与真实市场价格分布进行了对比。

首先我们可以从图中看到采用了 CE 模型的竞价策略出价的分布严重偏离了真实市场价格分布，绝大部分出价均偏高。然而从市场价格分布曲线中我们可以看到，真实的市场价格集中分布于 0 到 30 之间，在 50 到最高价格 300 的区间内逐渐下降。

而采用我们的 EU 模型的出价策略，其竞拍价格的分布同真实市场价格分布能够较好地吻合，代表着 EU 模型指导的竞拍能够恰好给出真实市场价附近的竞拍价格，也直接反映了我们提出的模型对于市场价格的真实感知，正如前文章节3.3.3所述。

另外，考虑到那些能够带来用户点击的广告展示才是有正回报的样本，我们统计了这些交易样本的市场价格分布并绘制在左图。从图中可以看到，我们提出的 EU 模型指导出价决策，其竞拍价格的分布更加符合带点击的广告展示的市场价格分布，这也一定程度上解释了我们模型能够获得较大利润提升的原因。

图3-4的右图绘制了不同模型驱动的出价策略其竞拍价格与真实市场价格差异的统计分布。我们可以直观看到，采用 CE 模型的竞价策略在竞拍时往往给出过高的竞拍价格，这可能会造成过多的广告投放预算浪费。而我们的 EU 模型，其竞拍出价非常接近真实市场价格，这也保证了模型能够保持较好的成本利用效率 (cost-effective)。

### 3.4.4 在线实验

我们的用户反馈预估模型于 2016 年 1 月在悠易广告公司<sup>1</sup>的在线广告平台上运行了两天。为了对比，我们上线了 4 个预估模型：EU, RR, CE 和 FM。其中前三个模型均为上文离线实验中介绍的模型，而 FM 是因式分解机<sup>[52]</sup>模型，是在点击率预估任务中被广泛采用的非线性模型。我们给所有的预估模型都设置了线性出价函数，并保持出价调整系数均一致以公平对比不同反馈预估模型的性能。所有费用的单位是人民币元 (CNY)。我们在 2016 年 1 月 25 日、26 日测试了 10 个广告活动，实时竞价交易的总量为 8,900 万，包括 3,300 万的广告展示，8,440 个用户点击和 1,403 元广告开销。预估模型在测试前 7 天的交易样本数据上进行训练，训练集包括 38,000 万广告展示，总计开销 52,500 元人民币。在测试过程中，每一个竞价请求将由实时竞价广告交易平台随机、等概率分配给不同的模型。同时我们给所有参与测试的模型设置了一样的预算限制。

图3-5展示了在线实验的模型结果。从结果中我们可以看到，EU 和 RR 模型均在另外两个传统模型基础上提高了利润和 ROI。特别地，EU 相比 FM 获得了两倍的 ROI 结果，而 RR 也获得了 50% 的 ROI 提升。EU 相比 CE 和 FM 分别将利润提高了 25.5% 和

<sup>1</sup>悠易互通广告公司官网：<https://www.yoyi.com.cn/>

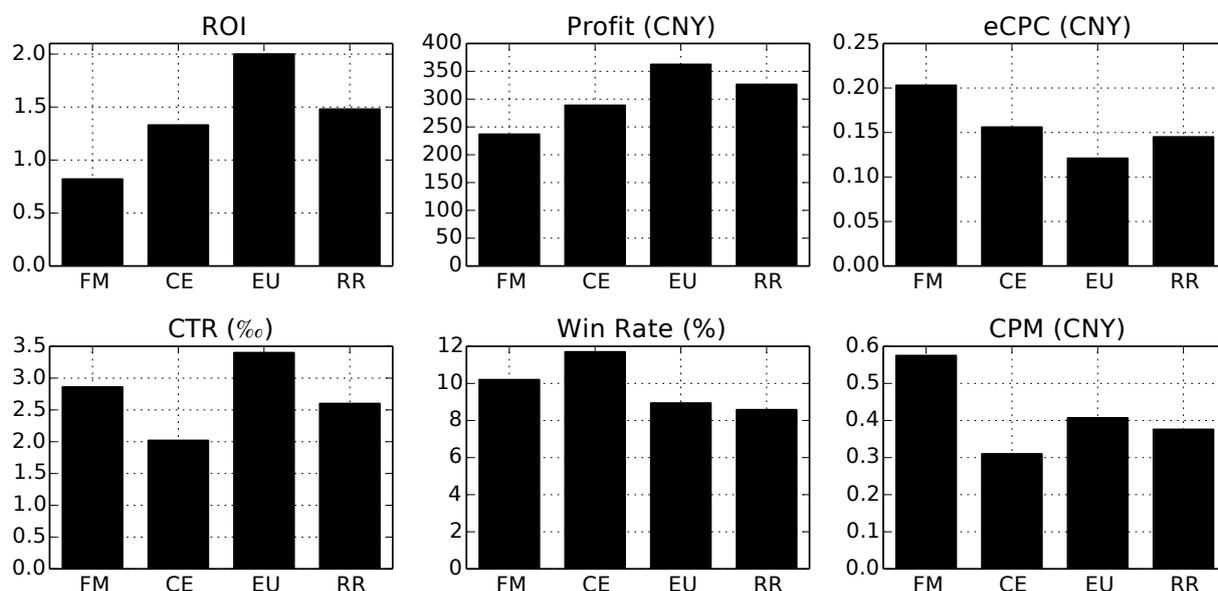


图 3-5 在线实验结果

Fig 3-5 Results of online experiments

53%。eCPC 指标的结果同 ROI 的趋势一致，即我们提出的模型 EU 和 RR 比另外两个传统模型均有较大的提升（更低的 eCPC）。对于点击率 CTR 指标，我们发现 EU 获得了最高的点击率性能。FM 比 CE 模型的点击率更高，这可能得益于非线性组合特征的利用<sup>[52]</sup>。EU 和 RR 模型赢得的广告交易相对更少，胜率不高，但二者的利润和 ROI 更高，这也从侧面反映出我们提出的模型对于成本效益的控制做得更为出色。

### 3.5 本章小结

在本章，我们针对用户反馈预估模型提出了全新的优化思路，创新性地将反馈预估模型的学习与后续应用场景的优化目标进行结合，并推导了模型学习过程的梯度。从数学推导结果来看，基于广告投放效益优化的用户反馈预估模型学习，本质上要求用户反馈预估对市场、竞价策略均有所感知。我们在离线数据集和在线广告平台上均做了全面的实验分析。实验结果表明我们的模型不仅在用户反馈分类问题上保持了优秀的表现，在广告主效益优化目标上相比传统模型获得了显著的提升。

表 3-4 广告活动利润提升效果对比 (相比基线 CE 模型)

Table 3-4 Campaign profit improvement over baseline CE

iPinYou	利润增益		ROI 增益	
	EU	RR	EU	RR
1458	7.10%	9.00%	233%	267%
2259	81.6%	99.3%	233%	472%
2261	26.3%	31.1%	44.4%	91.2%
2821	573%	615%	1334%	943%
2997	5.00%	0.700%	-3.60%	-11.4%
3358	1.70%	6.70%	77.1%	77.7%
3386	-1.20%	2.50%	20.6%	58.3%
3427	5.50%	8.70%	52.0%	175%
3476	4.20%	8.60%	16.0%	91.1%
YOYI	9.04%	0.600%	14.8%	2.11%
平均	+71.2%	+78.2%	+202%	+217%

表 3-5 其他离线实验指标效果对比

Table 3-5 Overall statistics in offline evaluation

iPinYou	CTR ( $\times 10^{-4}$ )				eCPC			
	SE	CE	EU	RR	SE	CE	EU	RR
1458	34	33	59	190	17	11	4.3	3.4
2259	3.3	3.6	3.7	5.8	303	235	172	136
2261	2.4	2.7	3.0	2.8	234	212	188	168
2821	5.5	5.9	4.8	7.0	116	137	105	112
2997	31	25	26	27	9.8	8.2	8.3	8.6
3358	51	41	69	61	18	19	12	12
3386	7.8	11	13	15	90	48	43	36
3427	7.2	25	29	72.8	98	25	17.3	10
3476	6.4	16	17	33.1	111	34	30	20
平均	16	18	25	<b>46</b>	110	81	64	<b>57</b>
YOYI	16	18	<b>26</b>	24	12.9	12.4	<b>11.3</b>	12

iPinYou	CPM				Win Rate			
	SE	CE	EU	RR	SE	CE	EU	RR
1458	57	37	25	65	0.22	0.24	0.13	.041
2259	100	84	64	78	0.89	0.63	0.44	0.24
2261	57	56	56	46	0.55	0.81	0.71	0.67
2821	63	80	50	78	0.12	0.63	0.48	0.45
2997	30	20	21	22	0.55	0.63	0.65	0.63
3358	92	77	80	70	0.11	0.20	0.11	0.13
3386	71	54	55	55	0.82	0.45	0.36	0.29
3427	70	60	49	75	0.75	0.26	0.22	.082
3476	71	55	50	65	0.49	0.31	0.31	0.15
平均	68	58	<b>50</b>	62	<b>0.50</b>	0.46	0.38	0.30
YOYI	<b>20</b>	23	29	30	<b>0.36</b>	0.30	0.22	0.22

## 第四章 竞价决策优化

第二章和第三章介绍了竞价形势预估和用户反馈预估，分别涉及到了实时竞价广告中成本预估和效益预估两大模块。本章将着眼于竞价策略的优化，即实时竞价广告中的决策部分，并提出整合效益预估、成本预估与竞价策略三部分的联合优化框架。

### 4.1 竞价策略优化

在上一章中我们介绍过，一般竞价策略优化由一系列优化模块构成。首先是用户反馈预估的模型优化，例如点击率预估，一般由机器学习的二分类模型进行建模与学习，如逻辑斯蒂回归模型<sup>[50, 51]</sup>、贝叶斯概率回归模型<sup>[1]</sup>、梯度提升决策树（gradient boosting regression trees, GBDT）模型<sup>[17]</sup>以及因式分解机模型<sup>[52]</sup>等。第二部分是市场竞争建模，即竞价形势预估模型。相关文献<sup>[24, 25]</sup>将其建模成关于市场价格分布的回归问题。基于前两部分的预测结果，第三部分竞价策略优化一般建模为带约束的优化问题<sup>[9, 18, 27, 63, 65]</sup>，并考虑其他的一些约束条件，例如广告投放的预算等等。

然而，这一系列独立的优化过程其实并非最优解决方案。根据贝叶斯决策理论<sup>[53]</sup>，用户反馈预估、竞价形势预估其实应该考虑到最终的竞价效益的目标。根据上一章的理论论证与实验分析，用户反馈概率预估结果在输出值域  $(0, 1)$  内并非拥有一致的意义，过低的预估结果可能导致广告主无法赢得广告交易；而过高的预估结果可能导致预算的浪费，无助于广告效益的提升。同时，不同广告活动的用户反馈价值又不尽相同。而用户反馈预估的结果又影响了实时竞价的结果，竞价形势预估的输出也对用户反馈预估模型的训练产生了较大的影响。

因此，市场价格和广告主之间的竞拍对广告活动的最终效果有着重要的影响。一方面，如果广告主的竞拍价格处于竞争激烈的价格区间，例如有较多的广告主倾向于在某价格附近进行出价，那么广告主赢得该实时交易的确信度（confidence）不会很高，这就要求用户反馈预估模型在这种情况下更加精细、精确地去进行预测。另一方面，竞价策略也在影响效益预估与成本预估模块的学习和预测。在相关文献<sup>[27]</sup>中，研究者证明了将点击率预估和市场价格预估进行结合的方法能提高竞价策略的效果。综上所述，一个自然的解决思路是将三种主要的挑战一起解决，通过一种综合优化方案对三个部分进行联合优化。

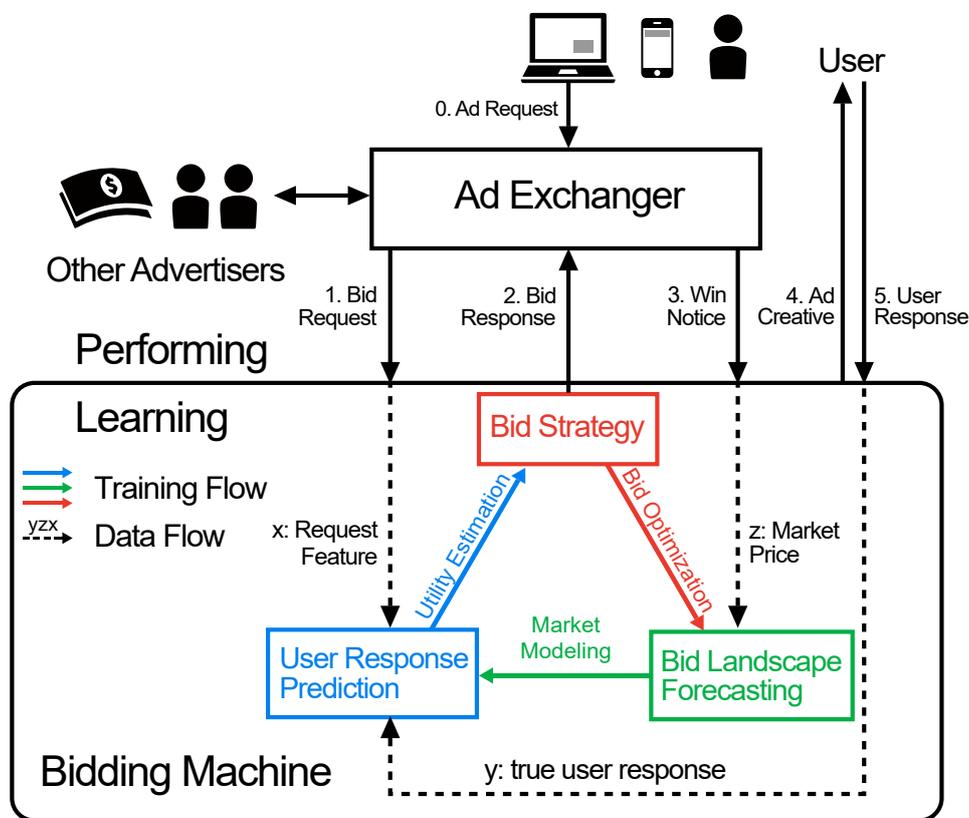


图 4-1 竞价机器 (Bidding Machine) 模型的联合优化框架图

Fig 4-1 The joint learning framework of bidding machine

### 4.1.1 竞价机器模型

在本章中，我们创新性地提出了一种联合优化框架，起名为竞价机器 (Bidding Machine, BM)，如图4-1所示。图中上半部分是实时竞价广告交易流程图，与第二章介绍的竞价交易机制相同，下半部分则是竞价机器模型的具体示意图。竞价机器模型同时考虑了用户反馈预估、竞价形势预估与竞价策略三部分，并采用同一种基于广告投放收益的优化目标函数。

在这个模型中，我们还考虑了广告主预算的限制，并提出了基于泛函优化方法的竞价策略优化理论，并详细推导了实时竞价广告交易中，最大化广告主利润的最优竞价策略函数的解析解。

竞价机器模型采用了一种“竞价学习” (learning-to-bid) 的迭代范式，它将最近的广告投放历史数据作为输入，然后更新预估模型参数；接着优化对应的竞价策略函数；最后在线上开展新的广告投放并收集实时用户反馈数据。完成一次完整的优化过程之后，竞价机器再次开始后续优化迭代。这个过程循环往复，就像是一台和市场进行不停交

互的自动优化机器，其目标是最大化广告投放的利润及效益，所以我们称之为竞价机器。

## 4.2 基于广告效益的联合优化

竞价机器模型的优化目标与第三章中的介绍的目标函数一致，即广告主的期望收益。但是在本章介绍的模型中，有三部分需要优化的参数或函数，即用户反馈预估模型的参数  $\theta$ ，竞价形势预估模型的参数  $\phi$  和竞价函数  $b(\cdot)$ 。竞价机器的优化问题则变成了得到最优的  $b(\cdot), \theta^*, \phi^*$  使得期望收益最大化，即

$$(b(\cdot), \theta^*, \phi^*) = \arg \max_{b, \theta, \phi} \int_{\mathbf{x}} R(\mathbf{x}, y; b, v, c, W) p_{\mathbf{x}}(\mathbf{x}) d\mathbf{x}, \quad (4-1)$$

我们采用了上一章介绍的期望收益模型 EU 的目标函数作为本章竞价机器模型的期望收益函数，即

$$\begin{aligned} R(b, \theta, \phi) &= \int_{\mathbf{x}} [vy - c(b(f_{\theta}(\mathbf{x})))] W_{\phi}(b(f_{\theta}(\mathbf{x}))) p_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \\ &= \sum_{(\mathbf{x}, y) \in D} [vy - c(b(f_{\theta}(\mathbf{x})))] W_{\phi}(b(f_{\theta}(\mathbf{x}))). \end{aligned} \quad (4-2)$$

### 4.2.1 针对用户反馈预估的效益优化

这部分优化同第三章介绍的优化方法一致，故不做详细介绍，仅将用户反馈预估模型 EU 的梯度推导结果再次列在下面：

$$\frac{\partial R_{\theta}^{\text{EU}}(\mathbf{x}, y)}{\partial \theta} = \overbrace{(vy - b(f_{\theta}(\mathbf{x})))}^{\text{竞价差异}} \cdot \overbrace{p_z(b(f_{\theta}(\mathbf{x})))}^{\text{市场敏感性}} \cdot \frac{\partial b(f_{\theta}(\mathbf{x}))}{\partial f_{\theta}(\mathbf{x})} \frac{\partial f_{\theta}(\mathbf{x})}{\partial \theta}. \quad (4-3)$$

### 4.2.2 针对竞价形势预估的效益优化

从第三章的推导及公式4-3可以看出，市场价格分布  $p_z(z)$  对于效益预估模块的优化有非常重要的作用。同时，正如章节3.3.3所述，市场价格分布函数  $p_z(z)$  有多种具体的实现形式。在本节，我们提出了一种基于机器学习方法的参数化模型来建模市场价格分布。另外需要注意的是，我们的模型学习目标是优化广告投放收益，而不是第二章介绍的真值数据的似然概率最大化。

我们依旧将公式 (4-2) 中的期望效益作为我们的优化目标函数, 然后针对市场价格分布预估模型的参数  $\phi$  进行求导获得偏导数结果如下

$$\frac{\partial R}{\partial \phi} = \frac{\partial}{\partial \phi} \left[ vy \int_0^b p_z(z) dz - \int_0^b zp_z(z) dz \right]. \quad (4-4)$$

接下来我们讨论几种市场价格分布函数的闭型 (closed form) 并推导每一种形式的优化梯度。

市场价格  $z \in [0, +\infty)$  是正数, 其概率密度函数是  $p_z(z)$ 。我们首先建模概率密度函数  $p_z(z, \mathbf{x}; \phi)$  并根据公式 3-1 推导出广告主出价  $b$  时的胜率函数  $W(b)$ 。我们的另一个目标是为胜率函数找到合适的闭型, 这就需要满足两个条件:

$$\begin{aligned} b \rightarrow 0^+, W &\rightarrow 0; \\ b \rightarrow +\infty, W &\rightarrow 1. \end{aligned} \quad (4-5)$$

我们从前文图 3-4 中我们可以发现, 市场价格分布具有长尾特性。所以我们提出了胜率函数建模的三种闭型。

**线性形式.** 我们假设市场价格分布在  $[0, \alpha(\mathbf{x}; \phi)]$  满足均一分布, 其中  $\alpha$  是市场价格的最大值, 由模型预测给出。其具体函数形式为  $\alpha(\mathbf{x}; \phi) = e^{\phi^T \mathbf{x}}$ , 输入的  $\mathbf{x}$  是竞价交易的特征向量。采用指数函数的原因是为了保证预估的市场价格的最值为正数。因此关于竞拍价格  $b$  的线性 (Linear) 形式胜率函数为

$$\begin{aligned} W(b, \mathbf{x}; \phi) &= \frac{b}{\alpha(\mathbf{x}; \phi)} = \frac{b}{e^{\phi^T \mathbf{x}}}, b \in [0, \alpha(\mathbf{x}; \phi)], \\ p(z, \mathbf{x}; \phi) &= \frac{\partial W(z, \mathbf{x}; \phi)}{\partial z} = \frac{1}{e^{\phi^T \mathbf{x}}}. \end{aligned} \quad (4-6)$$

其对应的期望收益目标 (如公式 (4-4) 所示) 的优化梯度  $\frac{\partial R}{\partial \phi}$  是

$$\begin{aligned} \frac{\partial R}{\partial \phi} &= \frac{\partial}{\partial \phi} \left[ vy \int_0^b p_z(z) dz - \int_0^b zp_z(z) dz \right] \\ &= \left( \frac{b^2}{2} - v y b \right) \frac{\mathbf{x}}{e^{\phi^T \mathbf{x}}}. \end{aligned} \quad (4-7)$$

如图 4-2 所示, 胜率函数  $W_z(b)$  随竞拍价格的增加而增长,  $\alpha$  控制了市场价格的最大值及市场价格分布的概率值, 每个样本的  $\alpha$  由模型进行预测。

**二次形式.** 我们依旧预测市场价格的最大值  $\alpha(\mathbf{x}; \phi)$ , 也就是市场价格与竞拍价格在  $[0, \alpha(\mathbf{x}; \phi)]$  之间。第二种闭型我们采用二次 (quadratic) 函数的形式来建模胜率函数, 即

$$\begin{aligned} W(b, \mathbf{x}; \phi) &= \frac{b}{\alpha(\mathbf{x}; \phi)} \left( 2 - \frac{b}{\alpha(\mathbf{x}; \phi)} \right), b \in [0, \alpha(\mathbf{x}; \phi)], \\ p(z, \mathbf{x}; \phi) &= \frac{\partial W(z, \mathbf{x}; \phi)}{\partial z} = -\frac{2z}{\alpha(\mathbf{x}; \phi)^2} + \frac{2}{\alpha(\mathbf{x}; \phi)}, \end{aligned} \quad (4-8)$$

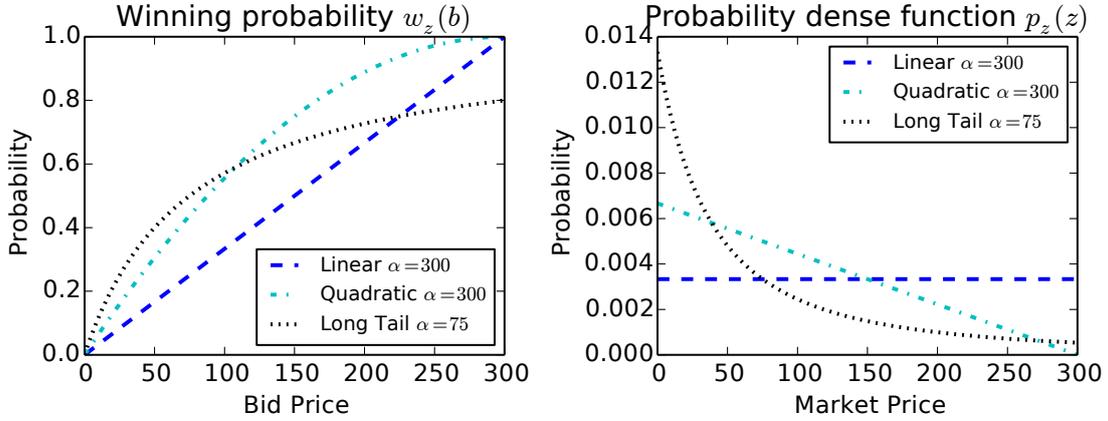


图 4-2 竞价形式与胜率函数的几种闭型示意图

Fig 4-2 An illustration of landscapes with different  $\alpha$  values

因为胜率满足  $W(b=0)=0, W(b=\alpha)=1$  的条件，如图4-2所示，并且在  $[0, \alpha]$  之间单调递增。

因此，期望收益目标函数关于竞价形势预估参数的梯度可以推导得到如下结果

$$\begin{aligned} \frac{\partial R}{\partial \phi} &= \frac{\partial}{\partial \phi} \left[ vy \int_0^b p_z(z) dz - \int_0^b z p_z(z) dz \right] \\ &= \left[ vy \left( \frac{2b^2}{\alpha(\mathbf{x}; \phi)^3} - \frac{2b}{\alpha(\mathbf{x}; \phi)^2} \right) - \left( \frac{4b^3}{3\alpha(\mathbf{x}; \phi)^3} - \frac{b^2}{\alpha(\mathbf{x}; \phi)^2} \right) \right] \\ &\quad \cdot \frac{\partial \alpha(\mathbf{x}; \phi)}{\partial \phi}. \end{aligned} \quad (4-9)$$

$\alpha(\mathbf{x}; \phi)$  是用回归模型拟合的映射函数，根据特征向量  $\mathbf{x}$  预测价格最值  $\alpha(\mathbf{x}; \phi)$ 。

**长尾形式.** 另外还有一种长尾 (long-tail) 形式的胜率函数被相关文献<sup>[9, 31]</sup> 所使用，即

$$W(b, \mathbf{x}; \phi) = \frac{b}{b + \alpha(\mathbf{x}; \phi)}, \quad (4-10)$$

我们依旧使用参数为  $\alpha$  的回归模型建模最优的市场价格最值参数。

长尾形式对应的市场价格概率密度函数是

$$p_z(z, \mathbf{x}; \phi) = \frac{\partial W(z, \mathbf{x}; \phi)}{\partial z} = \frac{e^{\phi^T \mathbf{x}}}{(z + e^{\phi^T \mathbf{x}})^2}. \quad (4-11)$$

因此我们能推导出其对应的模型参数梯度为

$$\begin{aligned}
\frac{\partial R}{\partial \phi} &= \frac{\partial}{\partial \phi} \left[ vy \int_0^b p_z(z) dz - \int_0^b zp_z(z) dz \right] \\
&= \frac{\partial}{\partial \phi} \frac{vyb}{b + e^{\phi^T \mathbf{x}}} - \frac{\partial}{\partial \phi} \int_0^b zp_z(z) dz \\
&= \mathbf{x} \frac{vybe^{\phi^T \mathbf{x}}}{(b + e^{\phi^T \mathbf{x}})^2} + \frac{\partial}{\partial \phi} \left[ \left( \frac{e^{\phi^T \mathbf{x}}}{e^{\phi^T \mathbf{x}} + b} + \ln(e^{\phi^T \mathbf{x}} + b) \right. \right. \\
&\quad \left. \left. - (\phi^T \mathbf{x} + 1) \right) e^{\phi^T \mathbf{x}} \right]. \tag{4-12}
\end{aligned}$$

正如图4-2所示，市场价格分布  $p_z(z)$  单调递减，因此当竞拍价格增大时，其对应的胜率函数取值趋向于 1。

### 4.2.3 竞价策略优化

接下来我们将针对竞价策略进行优化。优化过程将基于广告投放收益目标，并利用泛函优化方法推导最优化广告投放收益目标的最优竞价函数的解析解。有些相关工作<sup>[9]</sup>虽然也使用了泛函优化的方法，但是他的优化目标是用户点击数量而非广告投放收益，也没有考虑广告投放的成本，并且限于一价竞价（first-price auction）机制，没有在二价竞价机制下进行优化。

我们在附录B中证明了，在二价竞价机制下，最大化广告投放收益的最优竞价函数是线性竞价函数

$$\begin{aligned}
u(r) &= vr, \\
b(r) &= \gamma u(r) = \gamma vr. \tag{4-13}
\end{aligned}$$

其中  $r$  是用户反馈预估模型  $f_{\theta}(\mathbf{x})$  预估的用户反馈概率值， $v$  是每次用户正反馈行为能带来的收益， $\gamma$  则是调整系数。然后我们推导竞价函数中系数参数  $\gamma$  的最优解：

$$\begin{aligned}
\arg \max_{\gamma} T \int_r \int_0^{\gamma vr} (vr - z) p_z(z) dz \cdot p_r(r) dr \\
\text{s.t. } T \int_r \int_0^{\gamma vr} zp_z(z) dz \cdot p_r(r) dr = B. \tag{4-14}
\end{aligned}$$

$B$  为广告主总预算。

该最优化问题的拉格朗日函数是

$$T \int_r \int_0^{\gamma vr} [vr - (\lambda + 1)z] p_z(z) dz \cdot p_r(r) dr + \lambda B, \tag{4-15}$$

其中  $\lambda$  是拉格朗日系数。我们设定拉格朗日方程关于参数  $\gamma$  的导数为 0，则可以得到

$$\frac{\partial \mathcal{L}(\gamma, \lambda)}{\partial \gamma} = 0 \Rightarrow \gamma = \frac{1}{\lambda + 1}. \tag{4-16}$$

为了求解  $\lambda$ ，我们将公式 (4-16) 代入原问题的约束条件中可以得到

$$T \int_r \int_0^{\frac{vr}{1+\lambda}} z p_z(z) dz p_r(r) dr = B, \quad (4-17)$$

根据这个式子难以求解  $\lambda$  的解析解，除非我们能设定  $p_z(z)$  和  $p_r(r)$  的特殊形式。幸运的是，基于一些数值解法，我们能够求解出  $\lambda$  的值，因为方程 (4-17) 等号左边的部分关于  $\lambda$  单调递减。

从公式 (4-17) 可以看到，用户反馈预估结果  $r$  的概率密度  $p_r(r)$  直接影响了竞价函数 (4-16) 中的系数  $\lambda$  最优取值。这也就意味着，对于用户反馈预估模型  $f_{\theta}(\mathbf{x})$  的更新会导致预估结果输出的概率密度函数  $p_r(r)$  相应的产生变化，这也会导致竞价函数中  $\lambda$  的最优取值发生改变<sup>[9]</sup>。

#### 4.2.4 竞价机器优化算法

在确定了用户反馈预估、市场建模与市场价格预估及竞价策略优化的细节之后，我们给出竞价机器的详细学习算法，如算法4-1所示。

---

##### 算法 4-1 竞价机器的迭代优化算法

---

**Input:** 训练集合  $\{D_1, D_2, \dots, D_n\}$ ，每个训练集的总预算  $\{B_1, B_2, \dots, B_n\}$

**Output:** 最优反馈预估模型  $f_{\theta}()$ ，胜率函数  $W_{\phi}()$  和最优竞价策略函数  $b(\cdot)$

- 1: **for** 对每一个训练集合  $D_i$  **do**
  - 2:     (全优化)
  - 3:     **for** 对训练集合  $D_i$  中的每一个样本  $(\mathbf{x}, y, z)$  **do**
  - 4:         (双模块优化)
  - 5:         根据公式 (4-3) 和 (4-4) 计算对应模型的梯度
  - 6:         根据梯度计算结果优化模型参数
  - 7:     **end for**
  - 8:     根据公式 (4-17) 优化竞价策略函数  $b(\cdot)$
  - 9: **end for**
- 

结合图4-1分析算法。在获得广告请求后，我们的竞价机器会首先进行预测流程，即预估用户反馈概率和预测该交易的市场价格分布，然后通过竞价函数进行决策，给出最终竞拍价格。所有的预测结果会和交易结束后所获得的真值进行对齐 (alignment)。也就是说，如果广告主赢得了本次竞拍交易，那么广告展示后用户的真实反馈结果将用于监督用户反馈预估模型的优化，同时真实的市场价格也会被用来矫正市场价格分布的建

模与胜率预估模型的经验误差 (empirical error)。在这之后,我们会微调竞价策略函数来得到最优竞价函数,用于进一步最大化广告主的收益。这个离线的训练过程包含三个不同模型的交替训练和优化,即优化其中一个模型的同时固定另外两个模型的参数。若只对用户反馈预估模型与竞价形势预估模型进行联合优化,我们称之为“双模块优化”环节,而三个模块一起联合优化在称为“全优化”环节。我们将在实验中针对单模型优化、双优化与全优化进行对比分析。所有的模型参数更新将使用在线学习范式<sup>[51]</sup>。

竞价机器的迭代优化算法会将最近历史竞价数据作为输入用于三个不同模块的联合优化。关于在线优化算法的更新公式细节,我们将在附录A中进行详细推导。为了和第三中介绍的用户反馈预估模型进行公平对比,在实验中我们和反馈预估模型的参数更新方法保持一致,使用了随机梯度下降算法。

### 4.3 实验分析

在本章的实验环节,我们采用了同上一章第3.4节用户反馈预估模型的实验部分一致的数据集 (iPinYou 和 YOYI 数据集)、评估指标 (利润、投资回报率 ROI、点击率、有效点击费用 eCPC、每千次展示成本 CPM 和胜率),有所不同的是实验流程和对比模型。

#### 4.3.1 实验流程与对比模型

**实验流程.** 在上一章介绍的实验流程中,我们将原始广告展示数据作为离线竞价请求集合。经过特征工程等预处理以后,数据集包含一系列竞价记录三元组  $(\mathbf{x}, z, y)$ , 即竞价请求特征向量  $\mathbf{x}$ , 交易市场价格  $z$  和用户反馈标记  $y \in \{0, 1\}$ 。并且,我们和相关文献<sup>[9]</sup>采用了一样的实验流程:每一个竞价历史的特征向量根据历史上产生的时间顺序被当做离线竞价请求被发送给离线竞价引擎,离线竞价引擎会复现当时的竞价过程,即根据收到的竞价请求计算收益与成本并根据竞价策略进行出价。离线竞拍价格若大于数据样本中的市场价格,则记为赢得该离线广告展示,否则认为没有赢得该广告展示机会。第一阶段的实验是单模块优化与双模块优化的对比。在这个阶段,我们没有设置预算的限制,所有模型的竞价策略函数保持一致。

第二个阶段加入了广告主的总预算限制,用于对比不同竞价策略的预算分配效率及竞价效果。我们和相关文献<sup>[9, 65]</sup>采用了类似的实验流程,即设置了测试时的广告主预算分别为  $1/64, 1/32, 1/16, 1/8, 1/4, 1/2$  倍的训练集总预算额。这样对于不同的竞价策略函数来说,他们的竞价函数不同,但是总体预算约束是相同的。

**对比模型.** 在第一阶段实验中,我们对比分析了四个不同的模型。

**EU** 是我们第3.3.2节介绍的期望效益模型,其参数梯度如公式 (3-13) 所示。

**RR** 是我们第3.3.2节介绍的风险效益模型，其参数梯度如公式(3-14)所示。

**BM(MKT)** 是本章介绍的双模块优化方法所得到的用户反馈预估模型，在竞价时我们只需要使用用户反馈预估结果，但是其在训练过程中利用了市场价格建模部分的信息，所以我们在算法4-1中将其称为双模块优化的模型。

在第二阶段实验中，我们加入了广告主预算的限制。所以我们对比了四种不同的竞价策略。

**CELIN** 是上一章介绍的逻辑斯蒂回归模型 **CE** 作为用户反馈预估模型，并使用线性竞价策略函数的方法。这种竞价策略最简单，也是产业界最为普遍利用的竞价策略<sup>[9, 63]</sup>。其出价与用户反馈预估模型的输出成正比例关系，比例系数由人工调整决定。

**ORTB** 是用于优化点击数量的竞价策略，它与 **CELIN** 的不同之处在于使用了非线性竞价策略函数。具体可以参考相关文献<sup>[9]</sup>。

**PRUD** 是最近提出的基于市场价格预估模型的竞价策略，它根据预估的市场价格微调其出价，控制竞价策略的胜率与竞价效益。

**BM(FULL)** 是本章提出的全优化模型。根据算法4-1，每个模块都针对广告投放收益进行联合优化。

### 4.3.2 实验结果与对比分析

在本章节，我们分两个阶段进行实验结果展示与分析。第一阶段是双模块（用户反馈预估模型与竞价形势预估模型）优化方法与单模块（仅用户反馈预估模型）优化方法的效果对比分析，第二阶段是全优化（三个模块联合优化）的方法与传统竞价策略优化方法的对比分析。

**第一阶段实验.** 从算法4-1中可以发现，双模块优化方法综合了用户反馈预估模型与竞价形势预估模型的优化，后者的学习主要为了辅助前者在竞价过程中提高广告投放的收益。因此，在实验评估过程中，我们将双模块优化方法下收敛的用户反馈预估模型与其它用户反馈预估模型进行对比。

表4-1汇总了四个评估指标下的模型实验结果。从表格中我们可以看到，双模块优化下的用户反馈预估模型 **BM(MKT)** 在绝大部分指标下都获得了最好的效果。这个结果是比较容易理解的，因为双模块优化方法中两个不同模块的优化目标均为最大化广告投放的收益。从 **AUC** 指标的结果中我们可以看到，双模块优化方法中竞价形势预估的优化使得公式(3-7)中市场敏感性部分  $p_z(z)$  对分类准确率也有一定的正向贡献。

回顾第三章第3.4节中的实验结果，**EU** 和 **RR** 模型相比传统用户反馈预估方法，在利润与 **ROI** 方面已经有了巨大的提升。而本章介绍的双模块优化方法在这个基础上更

表 4-1 双模块优化与基线模型对比

Fig 4-1 Performance comparison between binary optimization model and baselines

		1458	2259	2261	2821	2997	3358	3386	3427	3476	平均
AUC	EU	.987	.674	.622	.608	.606	.970	.761	.976	.954	.795
	RR	.977	.691	.619	.639	.608	.980	.778	.960	.950	.800
	BM(MKT)	.981	.678	.647	.620	.603	.980	.788	.973	.955	<b>.803</b>
Profits ( $\times 10^7$ )	EU	3.91	.732	.797	.539	.147	2.42	1.58	3.05	3.25	1.82
	RR	3.98	.803	.827	.572	.141	2.54	1.64	3.14	3.39	1.89
	BM(MKT)	4.02	.766	.863	.669	.148	2.57	1.73	3.18	3.31	<b>1.91</b>
ROI	EU	19.2	.607	.582	.333	.679	9.26	1.46	5.30	4.02	4.60
	RR	24.3	1.03	.771	.247	.624	9.29	1.90	9.57	6.63	6.04
	BM(MKT)	31.7	.829	.692	.476	.733	8.83	1.08	9.70	5.40	<b>6.61</b>
eCPC	EU	4.27	172	187	104	8.33	11.4	42.5	17.3	30.0	64.3
	RR	3.39	136	167	112	8.61	11.4	36.1	10.3	19.7	<b>56.1</b>
	BM(MKT)	2.62	151	175	94.7	8.07	11.9	50.2	10.1	23.5	58.7

进一步，再次提高了利润与 ROI 指标的效果。

**第二阶段实验.** 第二阶段我们对比了四个模型，分别是 CELIN、ORTB、PRUD 和我们在本章提出的 BM(FULL)。另外在离线竞价实验过程中，我们为所有模型设置了 1/2, 1/4, 1/8, 1/16, 1/32 和 1/64 倍训练集总开销的预算设定，以对比不同模型在有限的预算条件下竞价效益性能的表现。

如图4-3所示，我们对比了所有模型在 iPinYou 所有子数据集上的综合实验结果。横轴表示不同的预算比例，纵轴是不同的性能指标数值。我们可以发现，在所有的实验设定和评估指标下，本章提出的竞价机器算法 BM(FULL) 获得了最好的实验结果。

从图中实验结果我们可以发现，由于 BM(FULL) 优化的是直接的广告投放收益，所以它在利润指标上获得了最好的效果。同时，当预算限制逐渐宽松（预算增多）时，所有的基线模型的效果会先提高后下降，这也说明这些基线模型并没有很好地关注竞价效益，以至于预算充足时有较多的预算被浪费。ROI 和 eCPC 的结果反映了成本效益。我们可以看到 BM(FULL) 的 ROI 指标在所有的实验条件下几乎没有下降，这也显示出我们提出的模型会动态调整出价策略，关注高收益的竞价机会，而非低效益的广告展示。

在表4-2中，我们列出了所有对比模型在 iPinYou 数据集上三种不同预算限制下的详细利润数值，表格数据清晰反映了上述描述的实验现象。

**模型收敛性能.** 我们在图4-4中绘制了竞价机器算法的三个不同模块在训练过程中的学习曲线。其中 ANLP 是第二章第2.5.2节介绍的负对数似然概率指标，用于衡量竞价形势预估模型的市场价格分布预测性能。从图中我们可以看出，三个不同的模块在评估各自性能的指标下均能快速收敛，同时三个模型又能联合优化共同的目标函数，使得广告投放收益获得大幅提高。

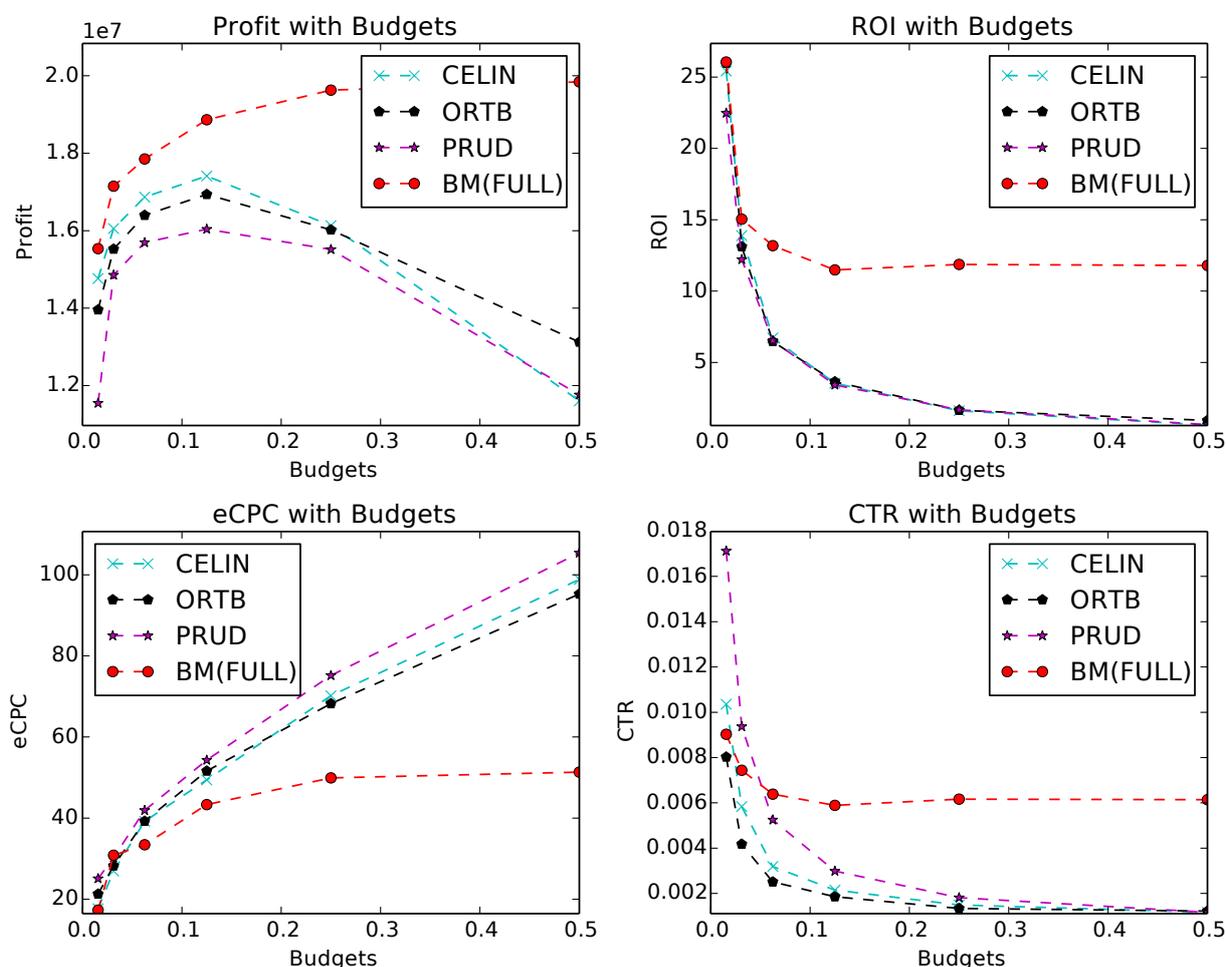


图 4-3 iPinYou 数据集上带预算 (budgets) 约束的竞价性能展示

Fig 4-3 Performance with budgets on iPinYou

### 4.3.3 在线实验

同第三章一样，我们也在在线平台上测试了本章新提出的算法。2017 年 4 月 1 日到 30 日，我们在 YOYI 在线广告平台上测试了三个模型，分别是 CELIN、EULIN 和 BM，其中 CELIN 和 EULIN 是第三章提出的 CE 模型和 EU 模型结合线性出价函数的方法，BM 则是本章提出的全优化竞价机器模型。整个在线实验包括 22,400 万的交易记录，2,300 万广告展示，16.8 万个点击以及总计 5 万人民币的累积广告开销。在线实验结果展示在图 4-5 中。

从图中结果我们可以看到，所有的竞价策略都获得了 100% 的利润收益。本章提出的联合优化模型则达到了利润与 ROI 指标上的最好效果。EULIN 模型获得了最高的

表 4-2 不同预算下的广告投放收益详情 ( $\times 10^6$ )

Fig 4-2 Achieved direct profit ( $\times 10^6$ ) with budgets

iPinYou	CELIN			ORTB			PRUD			BM(FULL)		
	1/64	1/8	1	1/64	1/8	1	1/64	1/8	1	1/64	1/8	1
1458	40.1	37.1	22.5	39.9	37.2	32.8	35.3	32.3	3.55	40.7	40.5	40.5
2259	2.47	5.67	-5.69	1.81	3.72	-2.55	4.00	5.20	.838	3.91	6.47	8.62
2261	1.93	4.48	.151	1.63	4.99	3.89	.872	3.47	2.46	1.97	5.31	9.13
2821	3.97	6.03	-13.8	3.69	5.47	-4.13	1.42	4.57	1.70	4.10	7.28	7.60
2997	.518	1.25	-1.03	.530	1.36	.151	.111	.665	.675	.670	1.39	1.49
3358	24.3	24.3	11.1	24.3	23.6	18.0	24.0	21.4	1.24	25.2	26.3	26.3
3386	7.83	13.3	2.43	6.67	12.3	10.1	7.10	11.5	3.71	8.30	16.4	18.4
3427	29.8	30.9	10.1	29.6	30.5	20.6	30.9	32.3	3.35	30.7	32.0	32.4
3476	21.7	33.3	15.7	17.3	33.1	22.8	19.8	29.6	2.82	23.3	33.6	34.8
平均	14.7	17.37	4.62	13.9	16.9	11.3	13.7	15.6	2.26	<b>15.4</b>	<b>18.8</b>	<b>19.8</b>

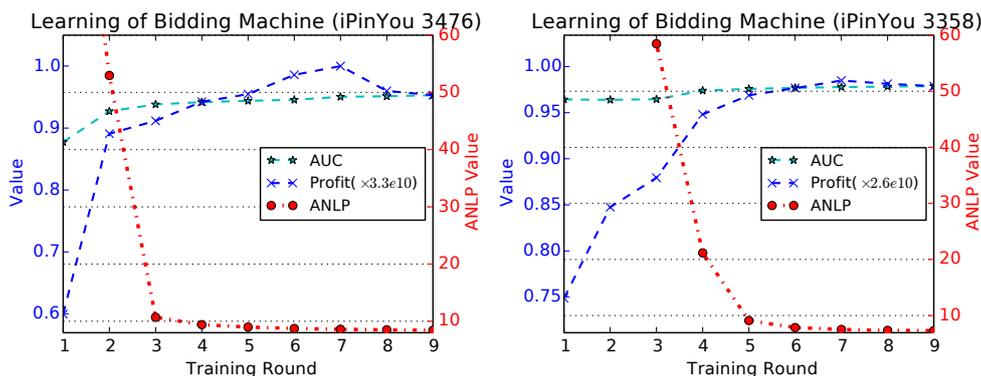


图 4-4 竞价机器模型不同模块的学习曲线

Fig 4-4 Bidding Machine learning performance

CTR 效果，但是其千次广告展示成本 CPM 偏高，这也导致了该模型在广告投放收益方面没有本章提出的竞价机器模型效果好。

### 4.4 本章小结

本章继续就广告投放收益目标进行优化，根据实时竞价广告的多模块共同作用的特点，创新性地提出了协同优化的综合优化框架：竞价机器。在竞价机器优化算法下，模型以在线更新的方式进行竞价操作与模型更新。在模型更新过程中，用户反馈预估、竞价形势预估与竞价策略函数优化三部分相辅相成，互相作用，共同优化统一的目标函

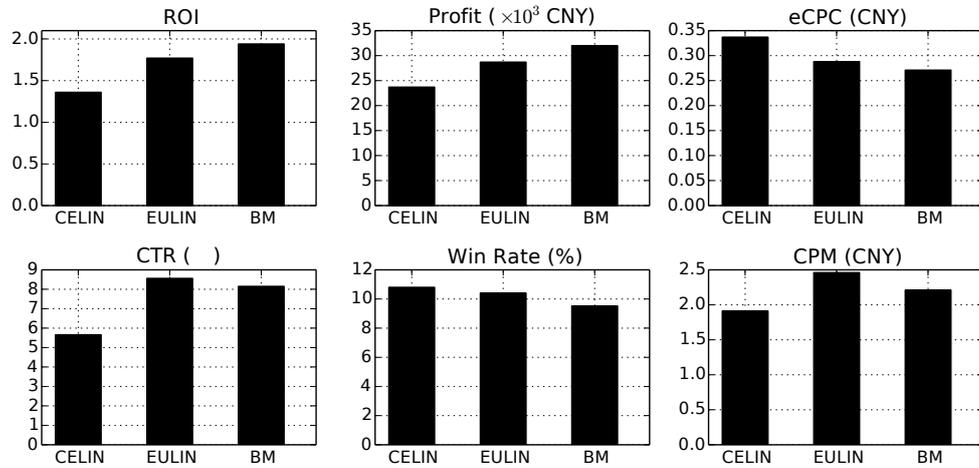


图 4-5 在线实验结果

Fig 4-5 Online results on YOYI platform in 2017

数，即广告投放收益。附录A章节将推导竞价机器模型的在线更新公式，B章节和C章节分别推导了最优竞价策略函数与博弈均衡分析。

在实验中，我们不仅在离线公开数据集上测试了我们模型的有效性，也在在线广告平台上验证了模型的效果。

## 第五章 终生用户建模

用户反馈预测，即建模用户对于给定物品的偏好，是在线服务，例如本文所涉及的实时竞价广告中非常重要的一部分。经过 20 年快速的发展，如今成熟的互联网平台上积累的用户从注册至今的行为序列已经十分丰富。每个用户不仅有较为固定的偏好，也有一些不断变化的兴趣。所以，为每个用户单独进行序列化建模，甚至是终生用户建模 (*lifelong user modeling*) 是一个很有挑战性的任务。目前存在的方法只能建模相对短期的用户行为，长序列尤其是终生行为序列建模为更好的用户建模提供了巨大的空间。此外，一个用户的行为可能与其之前整个行为历史中的很多次不同的行为有关，即，具有长期多跨度 (**multi-scale**) 的序列模式 (**sequential pattern**)。为了解决这些问题，我们在这篇论文中提出了一种基于记忆神经网络的方法，通过对每个用户个性化的序列模式记忆进行终生序列建模。该模型也利用了多层次多周期的更新机制来捕捉用户兴趣的多跨度序列模式，并实现了动态的更新以适应用户新近产生的行为。在三个大型真实数据集上的实验结果证明了我们的模型相对于当前最佳模型在用户反馈预测任务上有显著的提高。

### 5.1 用户行为建模

如今，准确地预测用户行为，例如点击或转化，已经成为了在线信息系统 (如搜索引擎<sup>[67]</sup>、推荐系统<sup>[15]</sup> 和计算广告<sup>[7]</sup>) 的核心部分。用户反馈预测的目标是估计用户对在线服务商提供的特定商品或内容给予预期行为的概率。估计的用户反馈概率可以引导在线系统的进一步决策，例如根据点击率对备选商品排序<sup>[15]</sup> 或根据用户反馈预估结果进行广告出价<sup>[9]</sup>，如前文第三章所示。

用户反馈预测的一个关键部分在于用户建模，即通过学习用户的历史行为和其他附加信息 (**side information**) 对用户进行画像。一般来说，用户的行为数据具有三个特征。第一，用户行为不仅反映了用户固有的 (**intrinsic**) 和多方面 (**multi-facet**) 的兴趣<sup>[68, 69]</sup>，也反映了用户兴趣的动态变化<sup>[70]</sup>。第二，如图 5-1 中左图所示，由于用户活跃度和注册时间的区别，不同用户的行为序列长度不同。第三，用户的行为历史中存在长期的依赖，如图 5-1 中右图所示，一些发生在很早的行为可能会影响用户后期的最终决定。此外，时序依赖也表现出多跨度 (**multi-scale**) 的序列模式，即不同用户的不同时间跨度上的行为依赖关系。

在线服务平台，包括在线广告平台经过近十年的快速发展，已经累积了十分丰富的

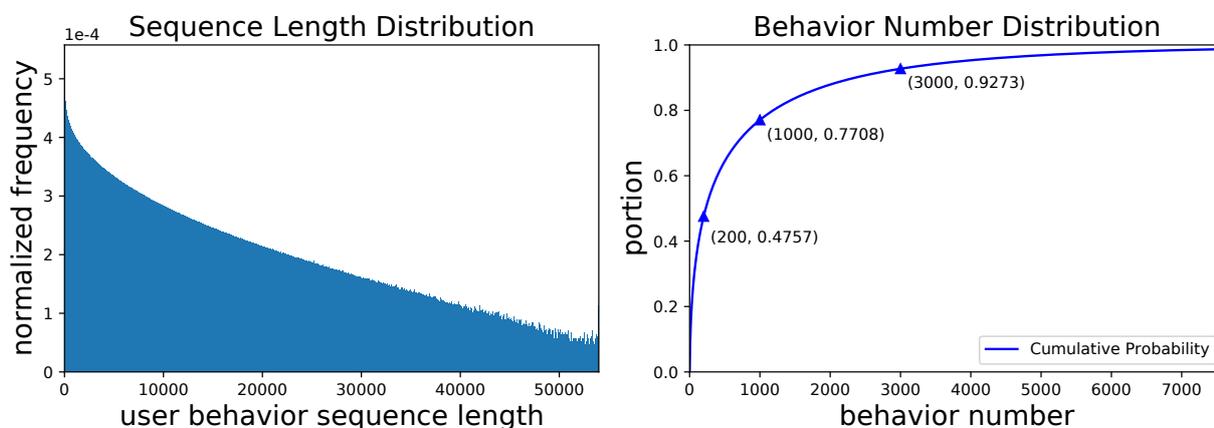


图 5-1 一在线电商平台的用户点击行为统计（2018 年 4 月至 9 月），左图：用户行为序列长度分布；右图：用户“加入购物车”与最终购买之间的行为数量统计

Fig 5-1 User behavior (click) statistics from a real-world e-commerce platform during April to September in 2018, Left: the distribution of the user sequence lengths; Right: the number of user behaviors between add-to-cart event and the final conversion.

用户行为序列数据。很多工作旨在解决基于历史行为的用户建模问题<sup>[19, 71]</sup>，尤其是通过序列建模 (sequential modeling) 的方式<sup>[43, 72]</sup>。早期的一些用户建模方法将用户行为历史聚合 (aggregate) 在一起以进行接下来的用户行为预测<sup>[69, 73]</sup>。然而，这些方法忽视了用户行为的时序动态特征<sup>[70]</sup>。用户反馈预测的时序建模即利用序列模式挖掘 (sequential pattern mining) 进行用户动态画像。其他一些工作<sup>[43, 72]</sup> 希望通过序列模式挖掘来解决时序行为预测问题。然而，这些序列建模工作只关注用户短期行为序列，即用户最近的若干次行为<sup>[43]</sup> 或近期一段时间的行为<sup>[72]</sup>，却抛弃了之前的行为序列。

考虑到人给人推荐物品的情况。人类一般会首先考虑他人固有的兴趣与品味<sup>[74]</sup> (tastes)，然后考虑其多面的兴趣<sup>[68, 69]</sup>，即对不同种类的商品表现出的多样的倾向性。此外，我们很自然地会想到结合一个人的长期<sup>[75]</sup> 和近期<sup>[76]</sup> 的经历来综合地推荐合适的商品。

为了解决这些问题，克服相关工作的弱点，我们提出了一个终生序列建模的框架并构建了一种新颖的多层级周期记忆神经网络 (Hierarchical Periodic Memory Network, HPMN) 来动态维护针对每个用户的兴趣与口味的记忆。具体来说，我们为每个用户建立了个性化的记忆存储，并通过模型学习到用户固有品味和多面兴趣的表征，压缩、存储在记忆中。模型维护了多层级的个性化记忆模块来保存关于特定用户长期行为的知识。HPMN 模型还根据用户新近的行为以不同的周期更新不同层级的记忆以捕捉多跨度的序列模式，并持续用户整个行为历史。我们在三个大型真实数据集上进行的实验证

明了我们提出的模型相对于此前最先进的模型有显著的提升。

本章提出的方法有如下三个主要贡献：

- 据我们所知，本章介绍的模型属于首次针对用户序列化建模提出终生用户建模框架，实现了统一、全面、个性化的用户画像。同时，模型面向超长用户行为序列数据，切实提高了用户反馈预测的准确率。
- 在终生序列建模框架下，本文提出了一种具有动态更新机制的记忆网络，利用用户历史上的行为序列数据和正在持续新增的行为数据进行动态用户建模。
- 我们进一步设计了多层次多周期的更新结构，以有效地挖掘并利用用户终生行为序列中多时间跨度的序列模式。

## 5.2 相关工作

### 5.2.1 序列化用户建模

用户建模，即捕捉用户可能的兴趣并得到每个用户的自适应表示，是用户反馈预估的关键部分<sup>[71, 77]</sup>。研究者们已经提出了很多用户建模方法，从隐语义方法<sup>[19, 73]</sup>到深度学习<sup>[15, 71]</sup>方法。这些模型将用户的所有历史行为聚合成一个整体，忽视了用户兴趣的时序性和动态变化。

如今，用户序列建模已然吸引了业界非常多的关注，因为用户行为序列中包含了更多关于用户兴趣的信息，尤其是兴趣变化的趋势与连续行为模式。在线系统中的序列化用户建模已经成为了研究的热点<sup>[43, 78, 79]</sup>。

从建模的角度看，用户序列建模方法主要分为三类。第一类是从时序矩阵分解的角度建模<sup>[70]</sup>，这种方法考虑到了用户偏好的变化，但启发性地对用户行为模式做了一些假设。第二类方法基于马尔科夫链<sup>[19, 80, 81]</sup>隐性地建模用户状态的变化并得到用户行为结果。第三类方法基于具有更强的特征提取能力的深度神经网络，如循环神经网络<sup>[72, 76, 79, 82-85]</sup>和卷积神经网络 (Convolutional Neural Network, CNN)，后者将用户行为历史作为图片进行建模<sup>[86, 87]</sup>。

然而，这些方法主要聚焦于短期用户建模，受限于用户最近期的行为。Zhang 等人在文献<sup>[74]</sup>中额外利用了一个静态的用户表示向量来建模用户固有的兴趣。Ying 等人<sup>[75]</sup>提出了用户行为序列特征上的多层注意力方法来建模用户长期兴趣。但是他们的模型只能获取较为简单的行为模式，没有考虑长期和多跨度的行为依赖。此外，已有工作几乎都没有考虑对用户终生行为序列建模，所以不能得到一个完整的、综合的用户画像。而且，就我们所知，这些工作都没有考虑长序列中多跨度的序列模式。

### 5.2.2 记忆神经网络

记忆神经网络<sup>[37, 88-91]</sup> (Memory Network), 又称记忆增强网络 (Memory-augmented Network), 最初被用于解决自然语言处理领域的一些问题, 它通过一个外部记忆模块显式地存储模型所提取出的信息。已有一些相关工作<sup>[92-95]</sup> 将记忆网络应用到推荐系统的任务中。但是这些方法直接运用了自然语言处理任务中的记忆网络结构, 没有考虑用户行为建模任务的实际情况。具体来说, 他们没有考虑对多跨度或长期序列依赖模式进行建模和记忆。

## 5.3 多层次周期记忆网络

在本节, 我们首先讨论用户建模与反馈预测任务的符号和基本描述, 在此基础上给出终生序列建模的定义并讨论该方案的一些特点。然后我们给出终生序列建模的整体框架, 包括多层次周期记忆网络 (HPMN) 的结构。表 5-1 总结了本章使用的符号。

### 5.3.1 符号与定义

在线系统中的数据可以用一个三元组集合表示  $\{(u, v, y)\}$ , 每个三元组包括用户  $u \in \mathcal{U}$ , 商品  $v \in \mathcal{V}$  和对应的用户行为标签

$$y = \begin{cases} 1, & u \text{ 和 } v \text{ 产生了交互;} \\ 0, & \text{反之。} \end{cases} \quad (5-1)$$

不失一般性, 我们同第三章一样, 把点击作为我们研究的用户反馈行为, 所以用户反馈预估的目标是预测用户  $u$  对商品  $v$  在特定时间的点击率, 并通过一个参数为  $\theta$  的函数  $f_{\theta}(\cdot)$  进行点击率预估。原始特征  $(\mathbf{u}, \mathbf{v}, \mathbf{c})$  分为三个部分。 $\mathbf{u} = (\bar{\mathbf{u}}, \{\mathbf{v}_i\}_{i=1}^T)$  是用户端特征信息, 包括用户的边际信息 (side information)  $\bar{\mathbf{u}}$ , 例如用户名等, 和一些列用户行为信息。 $\mathbf{v}$  是目标商品  $v$  的特征向量, 包括商品 ID 和一些附加信息,  $\mathbf{c}$  是预测请求的上下文特征, 例如网页 URL。

用户序列建模的目标是学习一个具有参数  $\Theta$  的函数  $g_{\Theta}(\cdot)$  来根据用户最近的  $s$  个行为为用户  $u$  生成一个综合的表示向量。

$$\mathbf{r} = g(\{\mathbf{v}_i\}_{i=1}^s; \Theta) \quad (5-2)$$

注意, 这个用户表示向量可能会发生变化, 因为用户不断地与在线系统进行交互并产生新的行为。很多用户序列建模工作设定了一个固定长度值  $s < T$  作为用户行为序列的最长长度, 如 Hidasi 等人的论文<sup>[72]</sup> 设定  $s = 5$ , 进行基于会话的推荐; Zhou 等人的论文<sup>[43]</sup> 设定  $s = 50$  来捕捉用户的近期兴趣。

在本章，我们定义用户反馈预测的最终任务是预测用户对于给定商品产生行为（如点击）的概率  $\hat{y}$

$$\hat{y} = \Pr(y|\mathbf{u}, \mathbf{v}, \mathbf{c}) = f(\mathbf{r}, \mathbf{v}, \mathbf{c}; \boldsymbol{\theta}). \quad (5-3)$$

表 5-1 本章所使用的符号标记与含义描述

Fig 5-1 Notations and descriptions in this chapter

符号标记	含义描述
$u, v$	目标用户和目标商品
$y, \hat{y}$	用户反馈真实标签与预估概率
$\mathbf{u}, \mathbf{v}, \mathbf{c}$	用户 $u$ 的特征, 商品 $v$ 的特征和上下文特征
$\bar{\mathbf{u}}$	用户的边际信息
$\mathbf{v}_i$	用户行为序列中第 $i$ 个交互过的商品
$\mathbf{r}$	查询获得的用户序列表示向量
$T, D$	用户行为序列总长度与 HPMN 总层数
$i, j$	用户行为索引与 HPMN 层索引 ( $i \in [1, T], j \in [1, D]$ ).
$\mathbf{m}_i^j$	第 $i$ 个行为节点第 $j$ 层维护的用户记忆槽内容
$w^j, t^j$	读取注意力权值与第 $j$ 层 HPMN 更新周期

### 5.3.2 终生序列建模

前文提到，大多数用户序列建模的已有工作聚焦于最近的  $s$  个行为，然而有时  $s \ll T$ ， $T$  是目标用户行为序列的完整长度。我们首先给出如下定义：

**定义 5.1 (终生序列建模 (LSM)).** 在用户反馈预测任务中指通过对用户终生行为历史进行序列模式挖掘从而实现连续的、动态的用户建模的过程。

我们总结了 LSM 的三个特点

- LSM 支持对用户行为模式的终生记忆。对于模型来说，要在实时预测与在线推断 (inference) 时，直接扫描并重新根据全部用户历史行为序列进行用户建模是不切实际的。所以 LSM 需要高效的信息保存机制。
- LSM 应对用户固有兴趣和变化的品味进行综合的建模以更好地预测用户行为。
- LSM 也应能持续动态地根据用户最新的行为进行更新。

依据以上的原则，我们提出了一种 LSM 框架对完整的、不断演化的用户行为历史进行建模，如图 5-2 所示。在这个框架中，我们为每个用户都建立了一个具备多个存储槽的个性化记忆模块。这个记忆模块会通过一个增量更新机制随着用户行为的演化进行更新 (如图 5-2 中步骤 A、B 所示)。

在线推断时，当一个用户发起访问请求，在线服务会转发这个请求的数据到记忆查询模块，包括目标用户和商品的信息。

每个用户请求都会触发一个查询过程，我们用目标商品的特征向量  $v$  作为查询向量，从每个用户对应的记忆池（memory pool）中取得相对应的综合用户表示。神经网络 HPMN 模型会用查询向量去读取该用户的个性化记忆池，而无需实时浏览整个用户历史序列，就可以快速得到实时用户表示向量。查询后，用户表示  $r$ ，商品向量  $v$  和上下文特征  $c$  会被一起用于后续的用户反馈预测中，我们将在章节 5.3.4 中描述这一过程。HPMN 模型的具体细节会在章节 5.3.3 中展示。

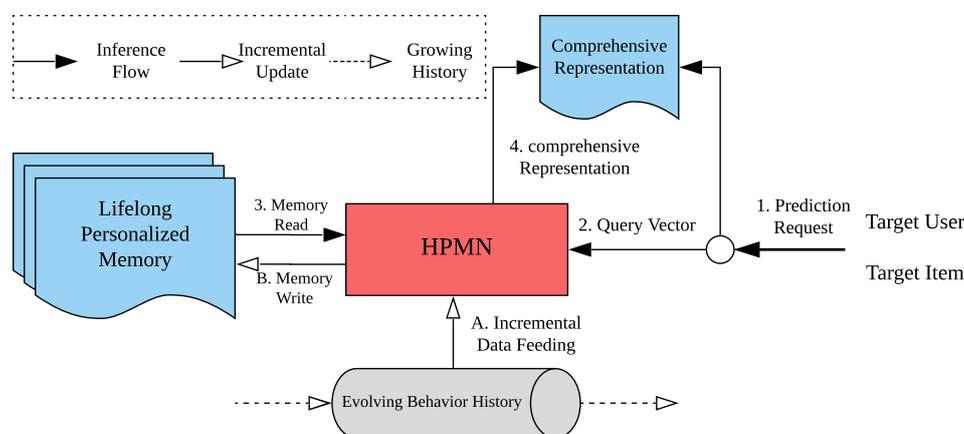


图 5-2 终生用户序列建模的示意图

Fig 5-2 The LSM framework

### 5.3.3 模型结构

在这一章，我们首先介绍 HPMN 模型的设计动机，再讨论具体的模型设计。总的来说，HPMN 模型的设计基于三方面的考虑。

- 如上所述，LSM 的主要目标是捕捉用户行为序列中隐藏的用户序列模式。已有很多工作<sup>[43, 76, 81]</sup>提出了一些序列模式挖掘方法。所以 HPMN 首先利用一个循环神经网络部件进行序列建模。
- 在用户终生行为中还存在长期序列依赖，即用户较晚的决定可能和很久之前的行为有关。我们会在实验部分展示一些例子。然而，传统的序列建模方法一般仅仅针对用户最近的行为进行建模；同时过于频繁地更新用户的状态向量会导致记忆饱和（memorization saturation）和信息丢失<sup>[96]</sup>。所以我们采用周期性的记忆更新机制来避免信息意外丢失。
- 行为依赖可能会跨越不同的时间跨度，即用户在行为历史中可能会对同一件或同一类商品在不同时刻表现出兴趣。所以我们需要多跨度的序列模式建模。HPMN 模型使用多层

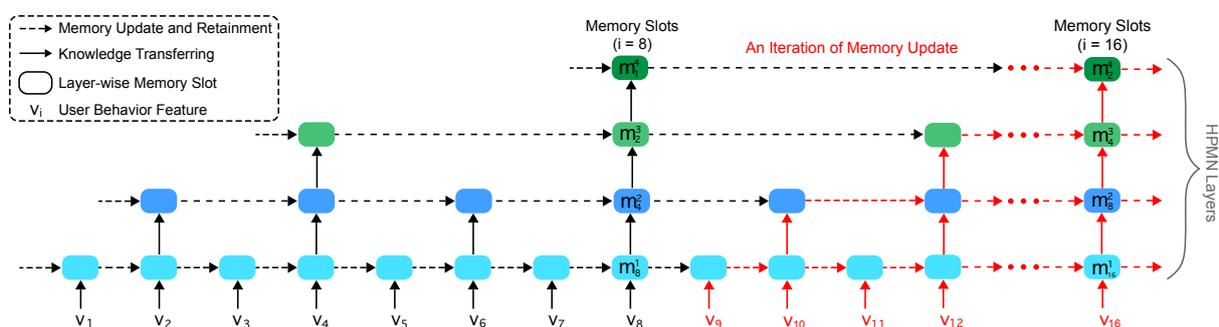


图 5-3 层级周期记忆网络 HPMN 的示意图；图中展示了 HPMN 维护的单个用户的记忆，共有 4 个 ( $D = 4$ ) 记忆槽。第  $j$  层的网络维护了更新周期为  $t^j$  的记忆，不同层的更新周期服从指数分布  $\{2^{j-1}\}_{j=1}^D$ ；红线表示一次完整的更新周期；虚线表示周期性的记忆和遗忘操作

Fig 5-3 The framework of HPMN model with four layers maintaining user memory in four ( $D = 4$ ) memory slots. The update period  $t^j$  of  $j$ -th layer follows an exponential sequence  $\{2^{j-1}\}_{j=1}^D$  as an example. The red part means the incremental updating mechanism; the dotted line means the periodic memorization and forgetting.

级多周期记忆模块来解决这个问题。

此外，因为个性化记忆存储了每个用户综合的、多面的兴趣，HPMN 模型利用记忆模块的协方差损失函数进行正则来保证存储知识的多样性。对于每次查询，模型通过注意力机制读取记忆，尝试将目标商品与用户多面兴趣进行相关性匹配。

下面我们会从四个方面描述模型的具体细节。我们会在章节 5.3.3.1 介绍记忆模块，接着在章节 5.3.3.2 描述周期增量更新机制。在章节 5.3.3.3 中我们会描述如何利用用户记忆，在章节 5.3.3.4 中描述协方差正则损失函数。

### 5.3.3.1 层级记忆模块

如图 5-2 所示，对每个用户  $u$ ，都存在一个用户特定的包含  $D$  个记忆槽  $\{m^j\}_{j=1}^D$  的记忆池  $\{m^j\}_{j=1}^D$ ，其中  $m^j \in \mathbb{R}^p$  是用户建模的实数向量表示。利用外部记忆的方法最初被用在自然语言处理领域<sup>[90, 97]</sup> 以对之前输入的文字段落中的信息进行更好地记忆。我们运用了外部记忆池来存储用户序列模式中的兴趣，它也支持动态更新以适应用户不断增长的行为序列。

如图 5-3 所示，总的来说，HPMN 是一个由  $D$  层组成的多层记忆网络。每层都维护特定的记忆槽中的内容  $m^j$ 。第  $j$  层在第  $i$  步 (即第  $i$  个用户行为) 的输出  $m_i^j$  不仅会被输入到记忆槽中存储已被后续读取与改写，也会被输入到当前步的下一层。

### 5.3.3.2 连续增量记忆更新

考虑到用户与商品之间交互的快速增长，模型预测时不可能实时浏览完整的用户行为历史。所以现有方法都只考虑短期的用户行为。我们的考虑是只维护最新的用户记忆，并实现一种实时增量更新方法。每当用户在第  $i$  步产生了一个行为，记忆槽会以如下方式更新

$$\mathbf{m}_i^j = \begin{cases} g^j(\mathbf{m}_i^{j-1}, \mathbf{m}_{i-1}^j) & \text{若 } i \bmod t^j = 0, \\ \mathbf{m}_{i-1}^j & \text{其他情况,} \end{cases} \quad (5-4)$$

其中  $j \in [1, D]$  且  $t^j$  是第  $j$  层的更新周期。在公式 (5-4) 中，每一层的记忆写入都是基于门控循环单元 (Gated Recurrent Unit, GRU) $g^j$ ，即

$$\begin{aligned} \mathbf{z}_i^j &= \sigma(\overline{\mathbf{W}}_z^j \mathbf{m}_i^{j-1} + \overline{\mathbf{U}}_z^j \mathbf{m}_{i-1}^j + \overline{\mathbf{b}}_z^j) \\ \mathbf{r}_i^j &= \sigma(\overline{\mathbf{W}}_r^j \mathbf{m}_i^{j-1} + \overline{\mathbf{U}}_r^j \mathbf{m}_{i-1}^j + \overline{\mathbf{b}}_r^j) \\ \mathbf{m}_i^j &= (1 - \mathbf{z}_i^j) \odot \mathbf{m}_{i-1}^j \\ &\quad + \mathbf{z}_i^j \odot \tanh(\overline{\mathbf{W}}_m^j \mathbf{m}_i^{j-1} + \overline{\mathbf{U}}_m^j (\mathbf{r}_i^j \odot \mathbf{m}_{i-1}^j) + \overline{\mathbf{b}}_m^j). \end{aligned} \quad (5-5)$$

每一层  $g^j$  的参数  $(\overline{\mathbf{W}}^j, \overline{\mathbf{U}}^j, \overline{\mathbf{b}}^j)$  不同。注意，和其他记忆网络如神经图灵机<sup>[91]</sup>相似，我们对记忆槽  $\mathbf{m}$  采用了软性写入方法，因为  $g^j$  函数中存在“擦除”向量  $\mathbf{z}^j$ 。第一层记忆  $\mathbf{m}_i^1$  根据用户商品交互的原始特征向量  $\mathbf{v}_i$  和上一步记忆向量  $\mathbf{m}_{i-1}^1$  进行更新。

此外，记忆更新是周期性的，每一层的记忆会根据当前时间步  $i$  和该层周期  $t^j$  更新。这里我们将每层的周期  $t^j$  设为超参数，具体设置在实验部分的表 5-3 中给出。通过运用这种周期更新机制，高层的更新频率会较低，以此可以实现两个目的。(1) 这避免了梯度爆炸 (gradient explosion) 或梯度弥散 (vanishing)<sup>[98]</sup>，从而可以对更长的序列进行建模；(2) 较高级维护的记忆可以比较低的层级更好地记忆更长期的行为依赖关系。每一层不同的更新周期可以使得模型能捕捉到多跨度的序列模式，我们将在章节 5.4.3 展示这一点。

相似的周期性更新方法在循环神经网络中已经有所应用<sup>[99]</sup>。但是他们只是简单的将循环单元的参数区分并分别进行更新。我们在此之上做出了两点提升：(1) 我们通过状态转换连接不同的网络层，使信息能够跨层传输；(2) 我们加入了外部记忆模块以保存固有和多跨度的序列模式以进行终生用户序列建模。

### 5.3.3.3 记忆读取

至此，模型已经实现了对用户固有和多跨度的时序动态特征的长期记忆，可以将行为依赖的固有属性和多跨度的模式与当前用户的行为预测相关联。此外，我们还通过与其他记忆神经网络相似的注意力机制<sup>[88, 89, 91]</sup>实现记忆读取。

模型在每次需要用户行为预测时进行记忆读取，以期获得用户综合表示向量。我们通过如下公式计算用户表示向量

$$\mathbf{r} = \sum_{j=1}^D w^j \cdot \mathbf{m}^j . \tag{5-6}$$

这里  $\mathbf{m}^j$  是长序列最后一步记忆的值，即  $i = T$ ， $T$  是用户行为记录中的最后一个。每个记忆槽的权值  $w^j$  代表每个记忆槽对于最后用户表示向量  $\mathbf{r}$  的贡献，以如下方式计算。

$$w^j = \frac{\exp(e^j)}{\sum_{k=1}^D \exp(e^k)} , \tag{5-7}$$

其中

$$e^j = E(\mathbf{m}^j, \mathbf{v}) \tag{5-8}$$

是一个计算查询向量  $\mathbf{v}$  和记忆槽  $\mathbf{m}^j$  相关度的能量模型。

注意能量函数  $E$  是一个非线性多层神经网络，使用 Relu 函数  $\text{Relu}(x) = \max(0, x)$  作为激活函数。

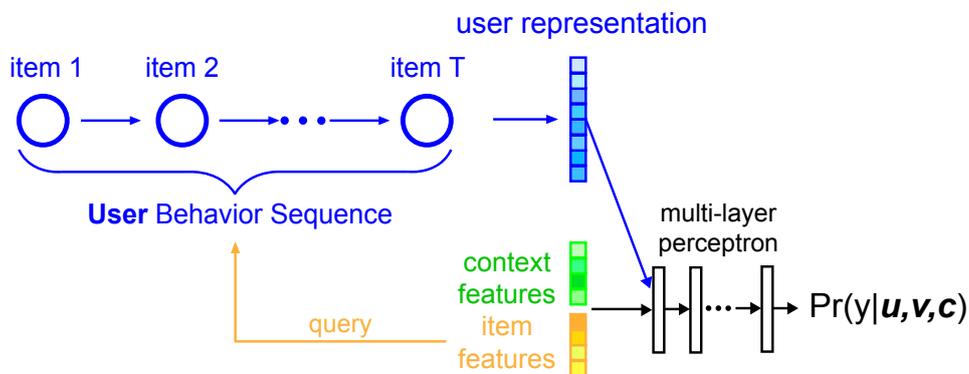


图 5-4 用户反馈预估的模型结构图

Fig 5-4 The overall user response prediction

### 5.3.3.4 协方差正则

如前面的章节所述，记忆模块捕捉了用户长序列多面兴趣模式。我们的模型用  $D$  个维度为  $p$  的记忆槽记忆用户的行为模式。我们希望不同的记忆能够存储用户兴趣的不同方面。然而，和神经图灵机<sup>[91]</sup>等模型不同，HPMN 在更新写入记忆内容时并没有用注意力机制进行内容筛选。所以为了提高记忆效率，我们在记忆模块写入机制的训练过程中加上协方差正则<sup>[100]</sup>。

具体来说，我们将记忆池内容的协方差矩阵定义为  $C$

$$C = \frac{1}{p}(\mathbf{M} - \overline{\mathbf{M}})(\mathbf{M} - \overline{\mathbf{M}})^\top \quad (5-9)$$

其中

$$\mathbf{M} = [m^1, \dots, m^j, \dots, m^D]^\top \quad (5-10)$$

是记忆矩阵， $\overline{\mathbf{M}}$  是  $\mathbf{M}$  每一行的均值矩阵， $p$  为每个记忆槽的维度。注意  $\overline{\mathbf{M}}$  和  $\mathbf{M}$  具有相同的形状。随后我们定义损失函数  $\mathcal{L}_c$  以实现正则

$$\mathcal{L}_c = \frac{1}{2}(\|\mathbf{C}\|_F^2 - \|\text{diag}(\mathbf{C})\|_2^2) \quad (5-11)$$

其中  $\|\cdot\|_F$  是矩阵的 Frobenius 范数。我们需要最小化不同记忆槽之间的协方差，即惩罚  $C$  的范数，使得记忆内容在写入时保持较好的差异性，从而能够建模用户不同方面的兴趣。

### 5.3.4 预测与损失函数

对每个用户反馈预测请求，我们根据公式 (5-2) 以及之前查询获得的综合用户表示向量  $\mathbf{r}$  进行预测。最终的用户行为估计结果可以如图 5-4 中所示，按以下方法计算

$$\hat{y} = f(\mathbf{r}, \mathbf{v}, \mathbf{c}; \Theta), \quad (5-12)$$

其中  $f$  是一个以 sigmoid 函数为激活函数的多层神经网络，其参数为  $\Theta$ 。

至于损失函数，我们采用端到端的训练方法并使用了交叉熵损失函数<sup>[23, 43, 71]</sup>  $\mathcal{L}_{ce}$  和协方差正则  $\mathcal{L}_c$ ，以及参数正则  $\mathcal{L}_r$ 。我们采用梯度下降作为优化算法。最终的损失函数为

$$\begin{aligned} \arg \min_{\Theta, \theta} &= \mathcal{L}_{ce} + \lambda \mathcal{L}_c + \mu \mathcal{L}_r \\ &= - \sum_{k=1}^N [y_k \log \hat{y}_k + (1 - y_k) \log(1 - \hat{y}_k)] \\ &\quad + \frac{1}{2} \lambda (\|\mathbf{C}\|_F^2 - \|\text{diag}(\mathbf{C})\|_2^2) + \frac{1}{2} \mu (\|\Theta\|_2^2 + \|\theta\|_2^2), \end{aligned} \quad (5-13)$$

其中  $\lambda$  和  $\mu$  是两个正则损失的权重系数， $\Theta = \{(\overline{\mathbf{W}}^j, \overline{\mathbf{U}}^j, \overline{\mathbf{b}}^j)\}_{j=1}^D$  是 HPMN 的参数集合， $N$  是训练集大小。

**模型讨论.** 我们提出了一种利用每个用户的个性化记忆进行终生用户序列建模的方法。用户记忆被周期性地更新，以捕捉一个用户长期多跨度的用户行为序列模式。当进行用

户反馈预测时，模型会根据目标商品向量对记忆模块进行，以预测用户对该商品的偏好。

注意，LSM 和相关文献<sup>[101]</sup>提出的终生机器学习 (lifelong machine learning, LML) 有很多关键的不同之处。首先，LSM 中模型维护的知识是针对特定用户的，而 LML 中知识是针对特定模型的；其次，LSM 聚焦于用户建模而 LML 的目标在于持续的多任务学习<sup>[102]</sup>；最后，LSM 中不断变化的是用户行为模式，而 LML 中数据和任务则在不断变化。

模型维护的高度压缩的记忆保证了我们模型在线推断时的时间复杂度是满足工业生产要求的。个性化的记忆从用户注册开始就被 HPMN 建立并维护以实现终生建模。对于每次预测，模型只需要查询保存的记忆，而不是向在其他工作中那样需要浏览全部的用户行为序列<sup>[43, 76]</sup>。同时，我们的模型相比于那些将用户行为聚合的模型，如隐语义模型<sup>[69, 70]</sup>，可以更好地实现用户序列建模。记忆更新的时间复杂度为  $O(DC)$ ，其中  $C$  是循环单元的计算时间。所有的矩阵操作可在 GPU 上并行完成。

HPMN 的参数可以像通常的方式那样被更新<sup>[15, 71]</sup>，即模型参数在一定条件下周期性地重新训练。记忆槽数量  $D$  是超参数，具体的数值需要基于实际情况确定随着用户序列的增长，用户记忆的大小也可以扩展。我们进行了相关实验并在章节 5.4.3 中讨论了记忆槽数量和模型效果的关系。

我们可以参照相关文献<sup>[96]</sup>的方法，在模型效果下降到一定程度时扩展记忆模块。如文献<sup>[96]</sup>所示，我们只需要增加一个更新周期更久的神经网络 HPMN 层在模型顶部，而不需要对所有参数重新训练。

## 5.4 实验分析

在这一章节，我们给出了实验设置的细节和实验结果。我们也针对我们模型的效果做了一些拓展及可视化分析实验。

我们通过三个研究问题来引导我们的实验和讨论

**RQ1** 终生行为建模是否有助于最后的用户反馈预测？

**RQ2** 在可比较的实验设置下，HPMN 是否取得了最好的效果？

**RQ3** HPMN 从用户行为序列中捕捉到了什么样的模式？它是否能够捕捉到长期，短期和多跨度的序列模式？

### 5.4.1 实验设置

在这一部分，我们给出了我们的实验设置，包括数据集描述，预处理方法，评价指标，实验流程和比较设置的讨论。

### 5.4.1.1 数据集

我们在三个真实数据集上运行对比了所有模型，三个数据集的统计信息在表 5-2 中给出。

**Amazon** <sup>[103]</sup> 数据集收集了来自亚马逊公司的用户对电商产品的浏览记录，包括商品评价和商品的属性信息。我们使用了其中的一个有关电子产品的子集，其中包括了用户从 1999 年 5 月到 2014 年 7 月的行为数据。我们将所有用户评价看作用户点击，并采样用户没有交互过的商品为负例，这种方法在相关工作<sup>[43, 71]</sup>中已被广泛采用。

**Taobao** <sup>[104]</sup> 包含了用户在淘宝平台上的行为数据。这个数据集包括几种用户行为类别：点击，购买，加购物车，喜欢与收藏商品。该数据集由接近一百万用户从 2017 年 11 月 25 日至 12 月 3 日的行为序列组成。

**XLong** 是从一个电商平台中超过两万个用户从 2018 年 4 月到 9 月的点击行为中采样出的。该数据集中包括的行为序列相比另两个数据集更长。目前的相关工作与公开数据集中尚没有包含如此长度行为序列的公开数据集。

**数据集特点** 这些数据集可被看作真实世界应用的一些典型例子。**Amazon** 数据集覆盖了很长一段时间 (大约 15 年) 的用户行为，但其中一些用户不太活跃，以至于产生的行为在这段时间中比较稀疏。在 **XLong** 数据集中，每个用户都有一个长度为 1000 的，时间跨度在半年的行为序列。对如此长的数据建模是终生序列建模的一个主要挑战。至于 **Taobao** 数据集虽然只包括了 9 天的数据，但其中的用户都有相当的活跃度，产生了较多的浏览行为。

**数据预处理** 为了模拟终生建模的环境，我们将用户行为按时间戳排序来构成用户终生行为序列。假设用户  $u$  有  $T$  个行为，我们用这个行为序列来预测用户对目标商品产生第  $(T + 1)$  个行为的概率。注意每个数据集中 50% 的目标商品都被替换成了另一个没有被目标用户点击的商品，以生成预测时的负样本。

**训练集与测试集划分** 我们根据需要预测的行为的时间戳划分数据集。如图 5-5 所示，我们在数据集覆盖的时间跨度中设置了一个切分时间。如果一个序列的最后一个待预测行为在切分时间前发生，我们就把它放入训练集，否则就将它放入测试集。以这种方法分出 70% 的数据作为训练集，30% 的数据作为测试集。

### 5.4.1.2 评估指标

我们用两个评价指标来评价用户反馈预测任务。第一个指标是 ROC 曲线下面积 (AUC)，该指标评估点击和未点击样本之间分类结果的排序表现。另一个指标为 **Log-**

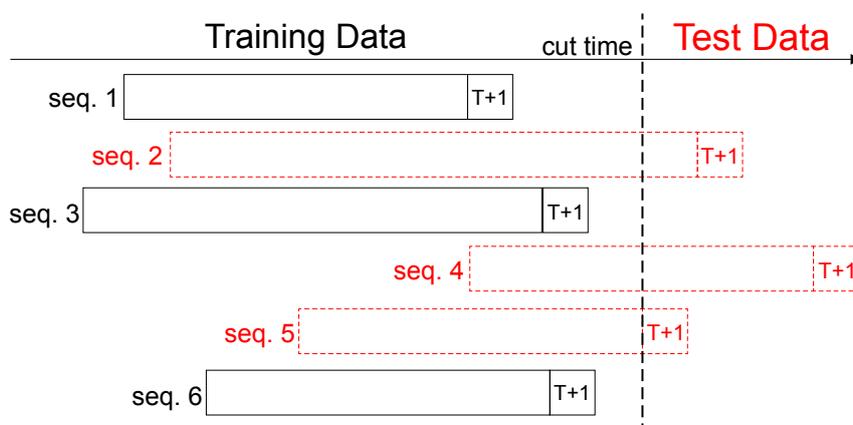


图 5-5 训练/测试数据集切分示意图

Fig 5-5 Illustration of data splitting.

表 5-2 数据集信息统计;  $T$  用户行为序列最大长度;  $s$ : 实验中近期行为序列的设定长度

Fig 5-2 The dataset statistics.  $T$ : length of the whole lifelong sequence (maximal length in the dataset).  $s$ : length of recent behavior sequence.

数据集	Amazon	Taobao	XLong
用户数	192,403	987,994	20,000
商品数	63,001	4,162,024	3,269,017
$s$	10	44	232
$T$	100	300	1,000

loss, 按以下公式计算

$$\text{Log-loss} = \sum_{k=1}^N \left[ -y_k \log \hat{y}_k - (1 - y_k) \log(1 - \hat{y}_k) \right]. \quad (5-14)$$

这里  $N$  是测试集中的样本数量  $\text{Log-loss}$  用来衡量模型整个数据集上的似然概率, 该指标在分类任务中被广泛采用<sup>[15, 16]</sup>。

### 5.4.1.3 实验流程

本章之前提到, 每个用户行为序列样本包含最多  $T$  个行为。因为我们的一些基线模型提出时只建模最近的短期行为序列, 所以我们先分割出  $s$  个用户行为作为短期序列数据以供基线模型使用 ( $s < T$ ), 具体设置在表 5-2 中展示。此外为了公平起见, 对于所有基线模型我们都进行了长度为  $T$  的序列建模实验。注意, 所有参与比较的模型都使用了一样的特征, 包括上下文特征和附加信息。

最后我们进行了**显著性测试**来验证我们模型相对其他模型的提升在统计学意义上的显著性。

#### 5.4.1.4 对比模型

为了展示我们模型的有效性，我们将它与三组共 8 个基线模型进行对比。第一组包含了基于聚合方法的模型，它们将用户行为聚合在一起进行建模和用户反馈预测，而没有考虑序列模式

**DNN** 是一个多层前馈深度神经网络模型，在最近的一些工作中被广泛用作基准模型<sup>[15, 71, 105]</sup>。我们参考<sup>[71]</sup>，用相加池化（pooling）操作来聚合所有的序列行为特征和其他特征的拼接作为用户表示向量。

**SVD++**<sup>[69]</sup> 是一个基于矩阵分解的模型，它结合了用户点击的商品和隐语义进行用户反馈预测。

第二组模型包含了短期序列建模方法，包括基于循环神经网络和卷积神经网络的方法和一个记忆网络模型。这些方法将近期的用户行为截短到特定长度  $s$ 。

**GRU4Rec**<sup>[76]</sup> 是一个循环神经网络的模型。这是第一篇利用循环神经网络单元来建模用户序列行为的工作。该工作最初提出时为了解决基于会话（session）的推荐问题。

**Caser**<sup>[86]</sup> 是一个基于卷积神经网络的模型，利用水平和竖直的卷积滤波器（convolutional filter）来捕捉不同尺度的行为模式。

**DIEN**<sup>[43]</sup> 是一个带有注意力机制的双层循环神经网络。它使用注意力值来控制第二层循环神经网络建模变化的用户兴趣。

**RUM**<sup>[93]</sup> 是一个神经网络模型，利用和自然语言处理领域中类似的外置记忆模块<sup>[91, 97]</sup>来存储用户的行为特征。我们按照<sup>[93]</sup>中的描述实现了特征层面的 RUM 模型（feature-level RUM）。

第三组包括一些长序列建模方法。但是，请注意本章提出的 HPMN 模型是由此提出用户终生序列建模的工作。

**LSTM**<sup>[37]</sup> 是第一个对长序列进行建模的模型，其记忆能力有限。

**SHAN**<sup>[75]</sup> 是一个多层注意力网络模型。它利用两个注意力层来分别建模用户的长期和短期序列。但是，该模型没有捕捉序列模式。

**HPMN** 是我们提出的模型，在章节 5.3 中有详细的描述。

我们先在较短（长度为  $s$ ）的数据上评估了第二组的模型，因为他们提出时就为了解决短期序列建模。然后我们在完整长度的数据上对比了所有的模型。

一些基于状态的用户建模方法<sup>[19, 80]</sup>在相关文献<sup>[86]</sup>中已经有了对比，所以我们只和当前最佳的用户序列化建模方法进行对比<sup>[86]</sup>。我们没有和另外的一些基于记忆模块的神经网络模型对比<sup>[92, 94]</sup>，因为他们的目标并不是用户序列建模。

在在线推断时，除了记忆网络（即 RUM 和 HPMN）以外的所有基线模型都需要载入全部的用户行为序列来进行用户建模，然而记忆网络只需要读取用户的个性化记忆内容

以进行进一步的预测。所以，考虑到线上运行，基于记忆网络的模型其空间效率更高。

我们的模型和其他记忆网络 (如 RUM) 的区别有两方面。(1) RUM 按照原文献<sup>[97]</sup> 中的描述实现记忆结构，该结构最初被用于自然语言处理领域，不一定适合用户反馈预测任务，因为用户生成的数据和语言中的句子具有较大差异。(2) 我们的模型利用周期更新的记忆模块和多层网络来捕捉多跨度的序列模式，RUM 没有考虑到这些特性。

#### 5.4.1.5 超参设置

超参数可以分为两组。第一组是训练超参数，包括学习率，正则损失项的权重系数。我们尝试了学习率  $\in \{1 \times 10^{-4}, 5 \times 10^{-3}, 1 \times 10^{-3}\}$  和正则权重  $\lambda, \mu \in \{1 \times 10^{-3}, 1 \times 10^{-4}, 1 \times 10^{-5}\}$ 。所有模型的训练批 (batch) 数据大小固定为 128。

表 5-3 不同数据集上的 HPMN 模型结构超参数详细设定

Fig 5-3 The HPMN structures on different datasets

数据集	记忆槽数量	不同层数及更新周期
Amazon	32	3 层每层周期: 1, 2, 4
Taobao	32	4 层每层周期: 1, 2, 4, 12
XLong	32	6 层每层周期: 1, 2, 4, 8, 16, 32

#### 5.4.2 实验结果与分析

在这一部分，我们在表 5-4 中给出了我们的实验结果，并从多个角度进行了分析。注意，之前章节 5.4.1.4 提到所有对比模型被分为三组，所以接下来的讨论我们也将分组展开。

**HPMN 和基线模型的对比。**从表 5-4 中可以看出，HPMN 相对其他基线模型有着明显的提高，取得了最佳的效果 (RQ2)。

第一组中基于聚合的模型，即 DNN 和 SVD++，比序列建模方法表现差。这表明用户行为数据中存在序列模式，对用户行为特征做简单的聚合会导致较差的效果。

和第二组中其他的序列建模方法相比，HPMN 的效果在不同长度的序列上都表现更好。GRU4Rec 在其原文献中主要为了解决短期的基于会话的推荐任务，所以它和 LSTM 有一样的问题，都会丢失与长期行为依赖相关的信息。虽然 DIEN 的注意力机制使它相对于单纯基于循环神经网络的 GRU4REC 有了明显的提升，但它忽略了多跨度的用户行为模式，我们会在下一节展示这一点。此外，DIEN 模型在在线推断时需要在整个用户序列上进行计算（因为需要针对每个用户历史行为做注意力计算），导致它在处理超长

表 5-4 性能对比结果；\* 代表显著性测试结果  $p$  值  $< 10^{-6}$ 。↑ 和 ↓ 表示在完整用户行为序列长度 ( $T$ ) 的数据集上模型性能优于/劣于基于短期行为序列 ( $s$ ) 的建模结果；性能最优的结果加粗显示，次优的结果则在下方划线

Fig 5-4 Performance Comparison. (\* indicates  $p$ -value  $< 10^{-6}$  in the significance test. ↑ and ↓ indicates the performance over lifelong sequences (with length  $T$ ) is better or worse than the same model over short sequences (with length  $s$ ). The second best performance of each metric is underlined.)

第 1 组	模型	序列长度	AUC			Log-loss		
			Amazon	Taobao	XLong	Amazon	Taobao	XLong
第 2 组	GRU4Rec	$s$	0.7669	0.8431	0.8716	0.5650	0.4867	0.4583
	Caser	$s$	0.7509	0.8260	0.8467	0.5795	0.5094	0.4955
	DIEN	$s$	0.7725	0.8914	<u>0.8725</u>	0.5604	0.4184	<u>0.4515</u>
	RUM	$s$	0.7434	0.8327	0.8512	0.5819	0.5400	0.4931
第 1 组	DNN	$T$	0.7546	0.7460	0.8152	0.6869	0.5681	0.5365
	SVD++	$T$	0.7155	0.8371	0.8008	0.6216	0.8371	1.7054
第 2 组	GRU4Rec	$T$	0.7760 ↑	0.8471 ↑	0.8702 ↓	0.5569 ↑	0.4827 ↑	0.4630 ↓
	Caser	$T$	0.7582 ↑	0.8745 ↑	0.8390 ↓	0.5704 ↑	0.4550 ↑	0.5050 ↓
	DIEN	$T$	<u>0.7770</u> ↑	0.8934 ↑	0.8716 ↓	<u>0.5564</u> ↑	<u>0.4155</u> ↑	0.4559 ↓
	RUM	$T$	0.7464 ↑	0.8370 ↑	0.8649 ↑	0.6301 ↓	0.4966 ↑	0.4620 ↑
第 3 组	LSTM	$T$	0.7765	0.8681	0.8686	0.5612	0.4603	0.4570
	SHAN	$T$	0.7763	0.8828	0.8369	0.5595	0.4318	0.5000
	HPMN	$T$	<b>0.7809*</b>	<b>0.9240*</b>	<b>0.8929*</b>	<b>0.5535*</b>	<b>0.3487*</b>	<b>0.4150*</b>

用户行为序列时缺乏效率。从基于 CNN 的模型 Caser 的结果表明，卷积操作可能并不适合于用户长序列建模任务。

至于 RUM 模型，尽管它也使用了外置记忆模块进行用户建模，但没有捕捉到序列模式，导致其结果较差。此外，该模型在原文献中主要针对不同于本章的评价指标进行优化 (预测准确率和召回率)<sup>[93]</sup>，所以可能在用户反馈预测任务中表现不佳。

通过对比 HPMN 和第三组中的模型 (LSTM 和 SHAN)，我们发现虽然基线模型都能聚焦于长期用户建模，但在超长用户行为序列建模中 HPMN 仍取得了更好的结果。其原因可能是因为 LSTM 的记忆容量有限，而 SHAN 则没有考虑用户行为的序列模式。

**有关终生序列建模的分析。**回顾前文，我们在短序列数据和终生序列数据上都评估了针对短期序列建模的模型。表 5-4 中列出了结果并用 ↑ (↓) 标出了后者相对于前者效果的提升 (下降)。从表中可以看出，几乎所有模型在建模 Amazon 和 Taobao 数据集时，长序列相比短序列均有提高。然而，在 XLong 数据集上，GRU4Rec, Caser 和 DIEN 的效果有所下降，不过基于记忆模块的模型 (如 RUM) 的效果有所提升。注意我们的 HPMN 模型取得了最好的效果。这些现象反映了终生序列建模有助于用户建模和提升用户行为预测的准确性 (RQ1)。但这也说明需要针对终生序列建模设计良好的模型结构，实验结果证

明我们的 HPMN 模型较好地完成了这一点。

**模型收敛性能.** 我们在图 5-6 中给出了 HPMN 模型在三个数据集上的学习曲线。如图所示，HPMN 收敛较快，在训练集上进行一轮迭代后，Log-loss 损失函数的值就达到了稳定的收敛点。

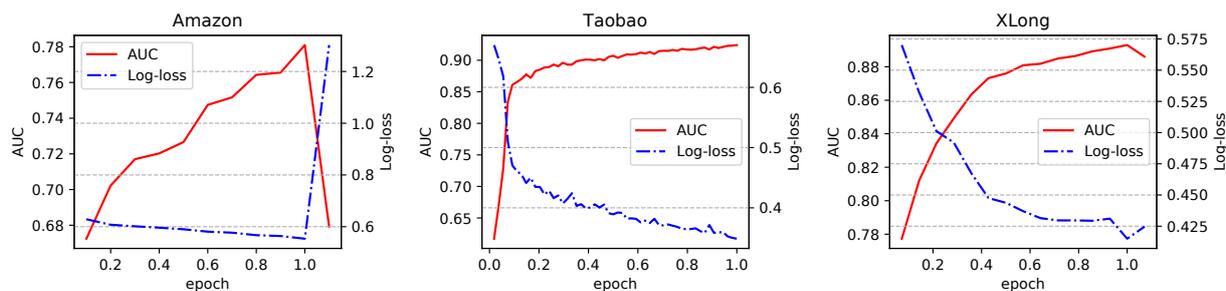


图 5-6 模型在不同数据集上的学习曲线，“epoch”表示训练数据集的一次完整迭代

Fig 5-6 The learning curve on three datasets. Here one epoch means the whole iteration over the training dataset.

### 5.4.3 可视化分析

在这一章节，我们进一步研究了 HPMN 在建模终生用户行为序列时所获取的模式 (RQ3) 以及模型的记忆能力。

**多跨度用户行为序列模式.** 在图 5-7 中，我们画出了 XLong 数据集中长度为  $T = 1000$  的用户行为序列中的三个真实样本。这三个序列分别反映了 HPMN 捕捉到的长期、短期和多跨度的序列模式。

在第一个例子中，目标商品是用户在预测时刻点击的“面霜”。我们在她的行为历史中发现，在她的第 31、33、37 个行为中对该类商品进行过点击，距离她最后的行为很远。当 HPMN 将目标商品作为查询语句读取记忆时，我们可以从注意力值热度图中看出第五个记忆槽的内容被分配了最高的注意力值，而其对应的第五层 HPMN 的更新周期相对较长。这说明了 HPMN 在其高层的记忆中捕捉到了长期的序列模式。

在第二个例子中，用户最后点击了桌子，且在最近点击了一些相似的商品 (桌子、柜子)。然而，在该用户行为序列的前部并没有发生对这些家具的点击。HPMN 第一层控制的第一个记忆槽被分配了最高的注意力值，表明较低的层可以更好地建模短期序列模式，其更新更加频繁，可以更好的捕捉用户的短期兴趣。

至于第三位用户，其对目标商品的点击包含长期依赖也包含短期依赖，她既在最近一段行为中，也在很久之前点击了相似的商品。当 HPMN 在  $(T + 1)$  时刻做推断时，第 2

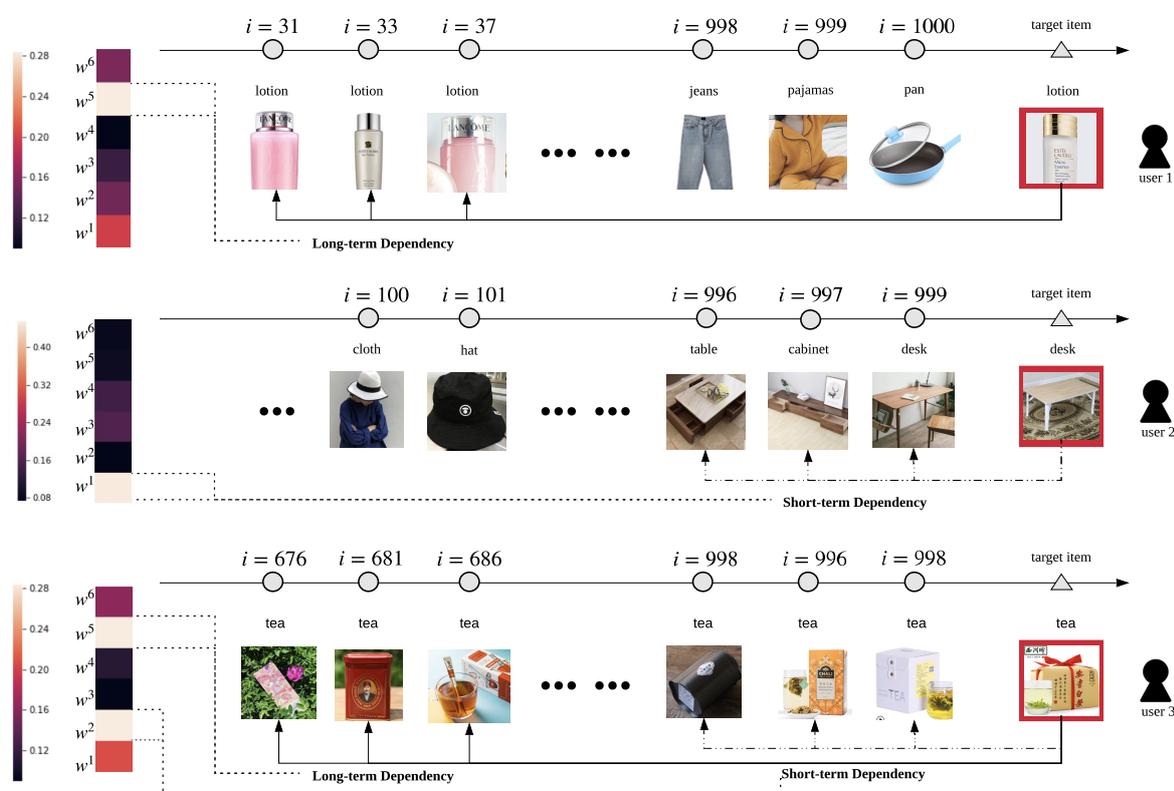


图 5-7 模型建模所得长短期及多跨度用户序列模式可视化

Fig 5-7 An illustration of long-term, short-term and multi-scale sequential patterns that are captured by HPMN.

和第 5 个记忆槽的内容被分配了较高的记忆里值，因为他们分别由第 2 和第 5 层 HPMN 控制操作，可以捕捉到短期和长期的依赖关系。所以这个例子展现了 HPMN 捕捉多跨度序列模式的能力。

**记忆能力.** 在图 5-8 中，我们展示了 XLong 数据集上不同总层数的 HPMN 模型的 AUC 表现。注意记忆槽的数量与 HPMN 的层数相同。从图中可以看出，当记忆槽数量小于 5 个，模型的预测表现随记忆槽数量增长快速上升。这表明模型用户建模需要更来的记忆模块来建模长序列的序列模式。且随着用户行为增加而增加记忆槽数量有助于捕捉终生序列模式。然而，当记忆槽数量大于 5，模型效果随着记忆增加有所下降。这表明模型能力根据用户行为长度而受到约束。这些结果可以提供了一些关于记忆模块进行扩展的指导，以及扩展 HPMN 模型的超参（例如层数等）的设计原则以实现面向不断增长的用户行为序列的终生序列建模，如章节 5.3.4 所述。

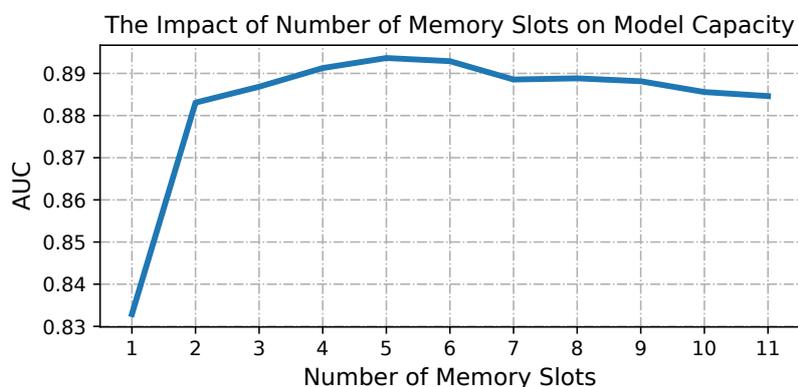


图 5-8 模型效果与层数及记忆槽数量的关系；第  $j$  层的更新周期数值服从层数相关的指数分布  $\{2^{j-1}\}_{j=1}^{11}$

Fig 5-8 The performance of HPMN with various memory numbers on XLong Dataset. The update period of each  $j$ -th layer follows exponential sequence  $\{2^{j-1}\}_{j=1}^{11}$

## 5.5 本章小结

在本章，我们提出了用于用户反馈预估的终生用户建模方法。为了完成这个目标，我们为每位用户构建了个性化的多层级记忆模块，用于存储其固有的和动态变化的用户兴趣，并由层级周期记忆神经网络进行维护。模型周期性地更新每个用户对应的记忆，以捕捉和存储多跨度的序列模式。用户的终生记忆会在之后预估行为概率时被读取。我们实验证明了终生行为序列建模的优势以及我们模型相对于其他已有模型包括目前最优模型的显著效果提升。

## 第六章 转化归因

我知道在广告上的投资有一半是无用的，但问题是我不知道是哪一半。

美国零售大王 John Wanamaker  
(1838-1922)

在在线广告系统中，用户在最终转化或购买广告商品之前，往往会浏览一系列的广告曝光 (impression)，例如展示广告、搜索广告或来自多个渠道的推荐商品。对于用户接触一次广告曝光的行为，我们称为广告的“触达” (touch)。对于广告主来说，准确估计用户在网上浏览过程中不同的广告触达对于用户最终转化行为的贡献值十分重要。这不仅有助于分析广告投放过程的投放效率，更能通过分析帮助指导后续广告投放的策略。在本章，我们将着眼于用户转化归因 (Conversion Attribution) 问题，针对现有解决方案的问题，创新性地提出基于深度学习与注意力机制的转化归因方法，并首次提出了一种离线评估框架，用以更高效地指导后续的广告投放活动。

### 6.1 多触达的转化归因

正如前文第一章所述，在线广告的一个优点在于广告主可以获得大量用户反馈，以衡量其广告活动的成功与否，并相应地对其进行优化。为了实现上述优化，计算广告的许多研究方向吸引了较多研究者的关注，并取得了一定的进步，包括用户定向<sup>[16, 51]</sup>，竞价策略<sup>[9, 23, 63]</sup> 和广告预算分配优化<sup>[26, 106, 107]</sup>。

如图 6-1 所示，通过在线广告，互联网用户可以在进行任何最终转化和交易之前，从多个渠道 (例如搜索引擎，社交媒体，移动平台) 接触一系列广告展示 (或称为广告曝光)，我们称之为广告触达 (touch point)。对于广告主而言，正确地将转化归因于每个广告触达点，即用户浏览历史中用户和广告内容之间的交互，是至关重要的。主要原因有以下三点：

1. 广告主应了解每个触达点对最终转化的贡献，以便做出合理的展示广告投放决策，包括实时竞价场景下的竞拍<sup>[108]</sup>。
2. 如果可以准确可靠地估算多次广告曝光的转化归因，则可以在广告主与代理平台 (以及广告技术提供商) 之间建立可量化的、基于贡献值的广告成本与收益分配方案。

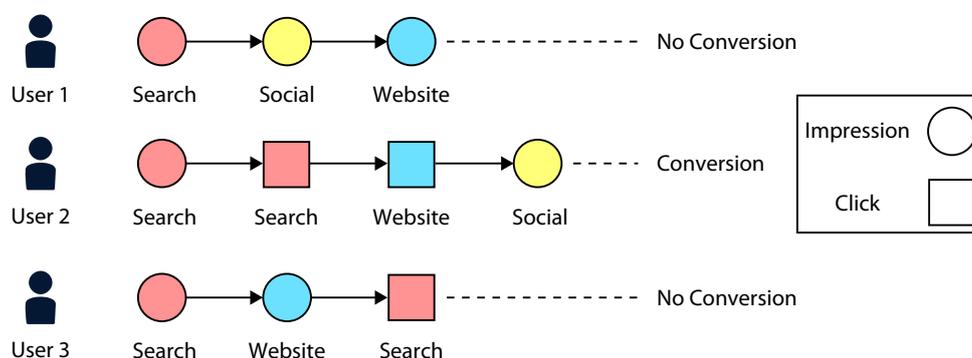


图 6-1 用户浏览与广告触达示意图

Fig 6-1 An illustration of user browsing and ad touch points

3. 根据广告渠道汇总的归因可以为广告主提供有价值的投放建议，以便为不同的广告投放渠道分配预算，从而在下一轮广告活动中以更低的成本获得更多的用户正反馈<sup>[109]</sup>。

目前，多数广告主仅使用一些基于规则的方法来解决转化归因问题，例如归因于第一次点击、最后一次点击等简单的机制<sup>[13]</sup>。具体来说，第一次（或最后一次）点击方法通常将转化完全归因于与广告内容的第一次（或最后一次）用户交互。虽然这些方法简单直接且易于部署，但显然它们缺乏足够的模式识别能力来支持更合理的预算分配<sup>[110]</sup>。Shao 等人<sup>[111]</sup>第一次提出了数据驱动的多触达归因模型，他们利用统计与机器学习的方法将广告触达对用户转化行为的贡献值分配到每一个接触点。之后，业界还出现了很多相关工作，包括使用一些基于分布函数假设的方法<sup>[64, 112]</sup>和加性激励过程<sup>[113-115]</sup> (additive exciting process) 等概率模型。

尽管这些方法的具有他们所声称的优点，但仍忽略了三个重要因素。

**序列模式建模。**上述方法都基于这一个假设：用户行为历史中的广告触达点独立地对用户转化作出正向贡献<sup>[111, 113]</sup>，而这点并不一定成立。实际上，用户浏览中的序列模式 (sequential patterns) 存在于每一位用户的上网行为中，例如用户在浏览了手机商品之后，可能会继续浏览手机配件，这种序列化的行为前后关系我们称之为序列模式。用户行为的序列模式在很多领域的行为预测或决策任务中都具有重大意义，例如推荐系统<sup>[19]</sup>，信息检索<sup>[116]</sup>和搜索广告<sup>[42]</sup>。如第五章所述，用户行为在用户建模与点击率预估等任务中承担着重要的角色。

**数据驱动的贡献值分配。**在这些方法中，贡献值是以启发式的方法分配到每次用户与广告的交互，而不是从数据中统计学习得到。例如，Ji 等人<sup>[114]</sup>提出，每个触达点根据当前转化的风险率 (hazard rate) 对最终转化以加性的方式作出贡献。这预先假设了用户经历越多广告曝光，转化的可能性就越高。然而如图 6-2 所示，在我们所使用的一个真

实数据集中，用户转化率并不随着用户行为序列长度的增加而提高。所以类似这种累积用户转化率提升的假设<sup>[113, 114, 117]</sup>可能会造成后续广告主更多地投放不合理的广告，并可能破坏用户使用线上服务时的体验。

**不同的转化前行为。**几乎所有的相关工作都忽视了多种用户浏览行为之间的不同之处。具体来说，他们仅仅基于浏览或点击行为进行归因分析，或者将所有行为等同视之。这些仅使用单一行为的方法无法达到很好的效果，因为用户通过不同的行为展示了明显不同的偏好，并可能在不同程度上引起最终的转化。

为了解决上述问题，我们在这篇论文中提出了一种双重注意力循环神经网络 (Dual-attention Recurrent Neural Network, DARNN) 模型以捕捉用户行为序列模式并学习最优的注意力分布作为归因。具体而言，我们的模型有两个学习目标。一方面，我们运用序列到序列 (sequence-to-sequence) 的结构<sup>[38]</sup>来建模广告曝光和用户的广告点击之间的关系，在这个过程中完成对点击行为的建模。另一方面，模型的最终输出是用户的转化率和模型对序列建模时的注意力值。注意力机制的优点在于它不仅可以提高模型预测的准确性，也自然地学习了转化归因在整个序列上的分布。此外，DARNN 不仅将注意力机制运用于原始触达点的特征，也运用在模型学习到的点击行为的隐藏层上，并且动态的结合两者来预测最终的转化率，即双重注意力机制。通过这种方法，模型不仅能够捕捉广告曝光层面的模式，也能够捕捉用户点击层面的模式。另外需要注意的是，双重注意力以及动态地将它们结合的方法都是从数据中统计学习出的，拥有较好的泛化性能。

此外，我们还提出了一种针对转化归因机制的离线评估框架。不同广告渠道获得的归因贡献值可以指导后续广告活动的预算分配，但目前没有相关工作讨论它们的归因方法的有效性<sup>[111-115, 117]</sup>，除非广告主花费大量预算进行在线模型测试，例如相关文献<sup>[64, 109]</sup>中的线上 A/B 测试。由于不同广告渠道的预算分配是广告主做出的重要决策之一，因此在线实验阶段之前能够评估归因方法的效果至关重要。

综上所述，本章介绍的工作具有四点创新性。

1. 我们构建了序列模式学习模型来建模用户行为序列；
2. 我们的模型在转化率预估任务中学习归因，而不是启发式地分配不同触达点的归因贡献值。
3. 我们的方法在多种不同类型的用户行为上进行了不同的、综合的归因分析。
4. 我们还提出了一个离线评估框架，通过广告预算分配和历史数据重放 (replay) 来衡量归因模型的效果。

本章后续展示的实验结果表明，我们的模型相比目前最优的方法在转化率预估问题上有显著提高。重放实验也显示我们的模型能获得目前最好的成本效益 (cost-effectiveness) 性能。

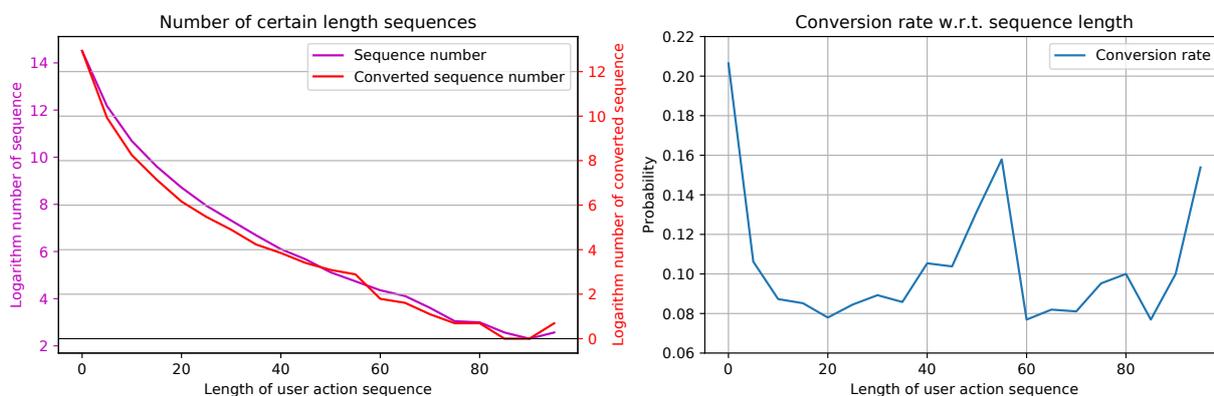


图 6-2 用户浏览行为序列长度统计及平均转化率随不同序列长度的变化（基于公开数据集）

Fig 6-2 The statistics of action sequence lengths and the conversion rates over Criteo dataset

## 6.2 相关工作

在线广告中，转化归因通常通过某些基于规则的方法计算，例如第一次点击或最后一次点击，然后根据获得的归因结果计算不同广告投放渠道的投资回报率 ROI，这可能会导致一些偏差<sup>[110]</sup>，因为实际上用户的转化一般由多次不同的广告触达共同作用影响。近年来，有很多基于多触达转化归因（multi-touch conversion attribution）的工作获得发表，用于对各种广告投放渠道中的一系列触达点进行归因<sup>[118, 119]</sup>。Shao 等人<sup>[111]</sup>第一次提出了数据驱动的多点归因模型，该方法基于用户浏览的广告使用分袋逻辑斯蒂回归（bagged logistic regression）模型预测用户转化率。其他一些工作主要基于简单分布假设的概率模型。Dalessandro 等人<sup>[112]</sup>提出归因贡献分配应该借鉴诸如经济学中的 Shapley 值的因果合作模型来计算。Xu 等人<sup>[64]</sup>认为广告曝光对用户最终转化具有不同的加性影响，并提出了一种基于提升的（lift-based）预测模型用于实时广告投放。但是，以上这些方法并没有将用户的所有接触点当做一个综合整体来考虑，忽视了时序和序列性因素<sup>[114]</sup>。并且，这些模型也没有很好地进行序列模式建模，而序列模式已被证明在用户建模中有显著效果<sup>[42]</sup>。

至于广告主和用户之间的多次交互，很多工作提出了用于用户建模的点过程模型。Yan 等人<sup>[120]</sup>提出了一个二维霍克斯过程（Hawkes process）模型来研究交易中卖家活动对获胜结果的影响。Xu 等人<sup>[115]</sup>提出了一个基于互激点过程（mutually exciting point process）的多点归因模型。该模型独立地将曝光和点击作为连续时间上的随机事件看待。这些点模型方法只考虑出现转化的情况，忽略了未转化的部分数据。在分析接触点的累积效应时，很多工作<sup>[64, 113, 114, 119]</sup>假设最终转化受用户浏览历史中触达点的加性影响。然而，这可能导致广告曝光呈现展示过多的趋势，可能会严重影响用户体验<sup>[121]</sup>。同

时，用户行为序列中的广告触达可能对最终转化起正面作用，也可能起负面作用，例如过多的展示会导致用户失去对展示广告内容的兴趣。所以，动态地为用户行为序列计算并分配归因贡献值是比较合理的方法。

另一类多点归因模型基于生存分析理论<sup>[113, 114, 117]</sup>，它将不同时候的用户转化行为事件建模为预测目标，并在考虑删失数据（如第2.2节所述）的同时估计特定时间转化事件发生的概率。然而，这些方法更多地关注单点预测，并且未考虑用户浏览历史中的序列模式。此外，其计算得出的归因贡献值主要是基于启发式的加性假设，这在实践中可能并不适用，例如相同内容的广告反而可能导致用户转化概率的降低。它们还假设了概率分布函数的具体形式，例如利用指数风险函数<sup>[113, 114]</sup>和 Weibull 分布<sup>[117]</sup>计算条件事件概率密度以使参数化模型可训练。然而，这样的分布形式假设可能会严重限制模型拟合真实数据的能力。

考虑到上述的限制，我们提出了一种双重注意力循环神经网络用户转化率预估和归因。注意力机制原本被用于机器翻译<sup>[38]</sup>任务，它通过一个序列到序列的模型计算下一个单词的概率，并在输出下一个单词时以不同的程度的注意力分布建模输入句子中的单词。在我们的场景下，注意力机制可以在动态地给特定用户的所有历史触达点分配归因贡献值。而用户行为中的序列模式则可以被循环神经网络很好地建模。

此外，一些工作讨论了在已有归因模型时的预算分配问题。Diemert 等人<sup>[108]</sup>提出了基于归因的实时广告系统出价策略，但这可能并不适用于更一般的在线广告场景。Geyik 等人<sup>[109]</sup>给出了一种基于归因得到的 ROI 值进行在线预算分配的方法。我们从该文献<sup>[109]</sup>借用了 ROI 的计算方法，并提出了一个基于多触达归因结果的离线测试框架。这是用于归因模型的第一个离线测试方法。

## 6.3 深度转化归因方法

在这一部分，我们首先描述多触达归因问题，然后提出利用双重注意力机制的序列行为建模方法，最后给出我们的对于预算分配的离线评估算法。

### 6.3.1 问题定义

不失一般性，我们从广告主的角度阐述问题。当一个用户  $u_i$  进行互联网活动（例如浏览网页，使用搜索引擎，使用社交媒体等）时，会和广告主投放的广告产生许多序列化的交互，即广告活动的触达点。对于一个共与广告主的广告产生了  $m_i$  个交互行为的用户  $u_i$ ，我们所观察到的是用户的浏览序列  $\{u_i, \{q_j\}_{j=1}^{m_i}, y_i, T_i\}_{i=1}^n$ 。其中  $y_i$  是表明用户是否转化的指示变量，如果产生了转化，则  $T_i$  是转化时间，否则  $T_i$  为空。每个触达

点  $q_j = (\mathbf{x}_j, z_j)$  由该接触点的特征向量  $\mathbf{x}_j$  和二值 (binary) 的行为类型  $z_j$  (曝光或点击) 标签构成。其中特征  $\mathbf{x}_j$  包括了用户和广告内容的特征信息, 例如用户 ID, 广告形式, 网站, 操作系统和浏览器信息, 也包括了广告投放的渠道标记  $c_j$  和交互发生的时间  $t_j$

我们的目标是建模用户序列模式以预估最终用户转化率并为用户浏览历史中的所有触达点  $\{q_j\}_{j=1}^{m_i}$  进行转化归因。一般来说, 归因越准确转化率预估的准确度就越高, 许多文献中都有类似的陈述<sup>[113, 114, 117]</sup>。在第6.3.2节, 我们给出了用于建模序列模式和预测最终转化率的循环神经网络结构。同时, 本章提出的模型还运用了序列到序列的神经网络结构建模用户点击模式, 让模型同时学习曝光和点击模式以更好地进行转化率预估。我们序列建模方法中的关键部分是双重注意力机制, 它将用户的两种行为 (曝光和点击) 纳入统一的综合框架并用于转化行为建模。我们将在第6.3.3节具体描述这一机制。最后, 模型从序列建模中获取的注意力值很自然地可以作为每个广告触达点对用户最终转化行为的归因贡献值。

有趣的是, 转化归因也有助于后续广告投放的预算分配<sup>[109]</sup>。在第6.3.4节, 我们会提出使用离线广告系列数据评估预算分配效果的方法。

### 6.3.2 序列建模

如图6-3所示, 我们利用循环神经网络进行用户序列建模, 类似的序列建模手段已广泛应用于信息检索系统<sup>[116, 122, 123]</sup>。请注意, 我们的方法的最终目标旨在进行最终转化率的预估, 而不是对每个触达点的点击率进行预测。

整个结构可以被分为三个部分: (1) 曝光层面的用户行为编码器; (2) 用于点击率预测的解码器; (3) 将以上两者结合以联合进行最终的转化率预估的双重注意力机制。我们会在本节阐明前两部分, 并在稍后讨论注意力机制。

**曝光层面的行为建模.** 对第  $i$  个用户行为序列  $\{u_i, \{q_j\}_{j=1}^{m_i}, y_i, T_i\}$ , 其中  $q_j = (\mathbf{x}_j, z_j)$ , 循环神经网络模型的输入特征序列是  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_{m_i})$ 。因为附带信息大多都是类别信息<sup>[105]</sup>, 我们先用一层嵌入层将稀疏的类别化 (categorical) 输入特征转化成稠密 (dense) 实数向量表示, 以用于后续训练。这种嵌入方法已被相关文献广泛采用<sup>[15, 21]</sup>。

然后将我们嵌入的特征向量输入到编码器 (encoder)  $f_e$  中, 表示为

$$\mathbf{h}_j = f_e(\mathbf{x}_j, \mathbf{h}_{j-1}), \quad (6-1)$$

其中  $\mathbf{h}$  是每一时间步的隐藏层向量。我们用标准的长短期记忆网络 (long short-term memory, LSTM)<sup>[37]</sup> 作为  $f_e$  的实现, 该结构在自然语言处理领域得到了广泛的应用, 并在第二章中被用于竞价形势预估模型的实现。

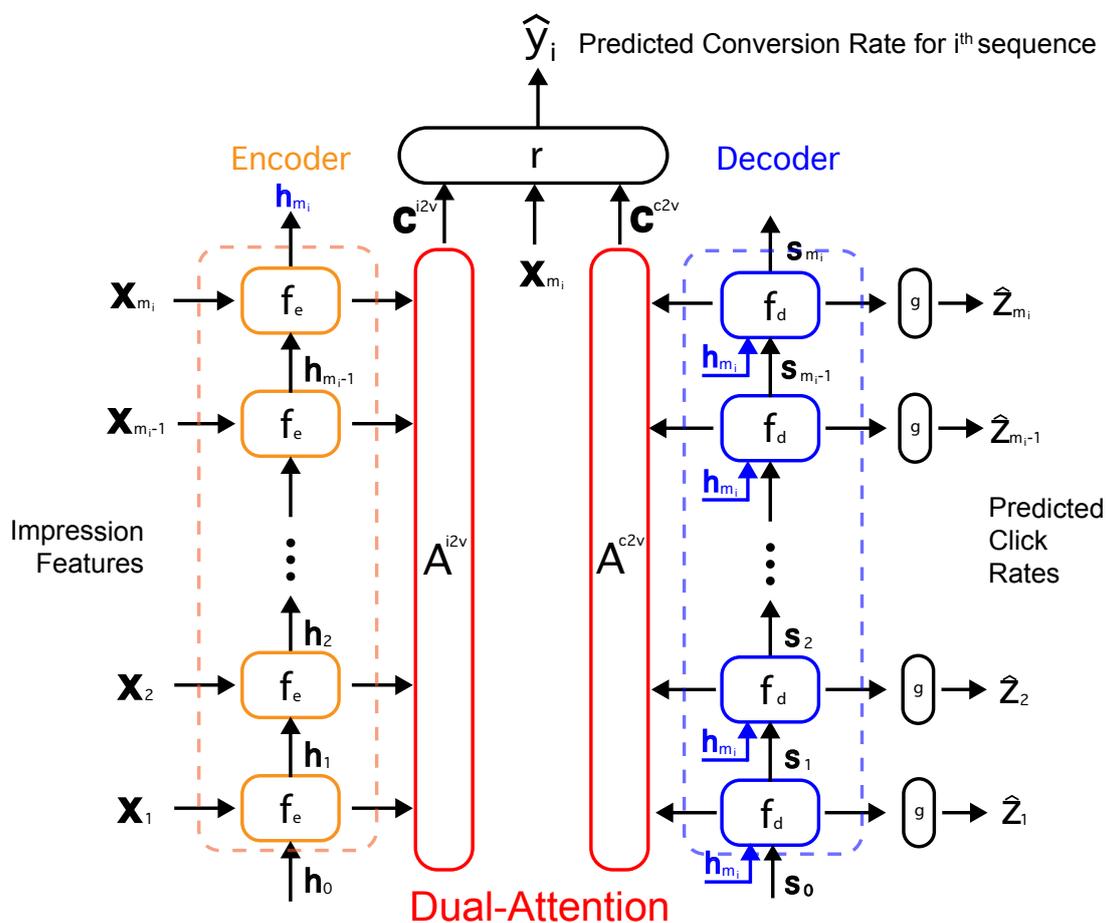


图 6-3 采用双注意力机制的序列建模

Fig 6-3 Sequential modeling with dual-attention conversion attribution

**点击层面的行为建模.** 在这一部分，我们的目标是建模广告展示时用户的点击行为。

在序列到序列模型中，解码器（decoder）通过将联合概率分解成有序的条件概率得到用户点击的联合概率分布  $\mathbf{z}$ 。

$$p(\mathbf{z}) = \prod_{j=1}^{m_i} p(z_j = 1 \mid \{z_1, \dots, z_{j-1}\}, \mathbf{x}). \quad (6-2)$$

注意其中  $\mathbf{z} = (z_1, \dots, z_{m_i})$  且  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_{m_i})$ 。通过这个解码器，每个条件概率可以被建模成

$$\hat{z}_j = p(z_j = 1 \mid \{z_1, \dots, z_{j-1}\}, \mathbf{x}) = g(z_{j-1}, \mathbf{s}_j). \quad (6-3)$$

这里  $g$  是一个以 Sigmoid 函数  $\text{sigmoid}(x) = \frac{1}{1+e^{-x}}$  为激活函数的多层全连接感知机 (Perceptron)<sup>[124]</sup>，输出为  $z_j = 1$  的概率。 $\mathbf{s}_j$  是点击层面的第  $j$  个接触点的隐藏层向

量。该值通过以下公式计算

$$\mathbf{s}_j = f_d(\mathbf{s}_{j-1}, z_{j-1}, \mathbf{h}_{m_i}), \quad (6-4)$$

其中  $f_d$  是非线性循环神经网络解码器函数，用于对用户点击模式进行建模。我们同样采用 LSTM 作为它的实现。在预测每一个隐层向量  $\mathbf{s}_j$  时，解码器都会将编码器最后一个隐层向量  $\mathbf{h}_{m_i}$  作为输入。我们的第一个损失函数就基于序列预测得到的点击率进行计算：

$$L^c = \sum_{i=1}^n \sum_{j=1}^{m_i} -z_j \log \hat{z}_j - (1 - z_j) \log(1 - \hat{z}_j). \quad (6-5)$$

在本章介绍的模型中，我们利用序列到序列进行点击率预测有两个原因。第一，这种多行为预测的任务某种程度上与多任务学习的思想相似，可以减轻数据稀疏性问题并为用户行为特征建立共同的基础隐空间向量表示。众所周知，用户一般遵循这样的行为模式：他们可能会在广告曝光后进行点击，而一系列的连续广告投放可能会驱使他们产生转化购买行为，这种“曝光-点击-转化”的行为模式导致了数据稀疏性问题<sup>[125]</sup>。具体而言，点击远不如曝光频繁而转化则更为稀少。所以有必要采用一种方法来解决数据稀疏性问题。直觉上，我们的模型是利用点击行为的信号来帮助提高稀疏的转化行为的预估精确度。另一个原因是为了获得点击层面的归因，在我们后续讨论的实验结果中，我们可以发现点击行为往往能获得比曝光更高的归因贡献值。

### 6.3.3 归因学习

我们最终的目标是建模用户序列模式并预测转化率，所以模型最后的输出通过以下公式计算：

$$\begin{aligned} \mathbf{c}^{i2v} &= A^{i2v}(\mathbf{h}_1, \dots, \mathbf{h}_j, \dots, \mathbf{h}_{m_i}), \\ \mathbf{c}^{c2v} &= A^{c2v}(\mathbf{s}_1, \dots, \mathbf{s}_j, \dots, \mathbf{s}_{m_i}), \\ \hat{y}_i &= p(y = 1 | \mathbf{x}, \mathbf{z}) = r(\mathbf{x}_{m_i}, \mathbf{c}^{i2v}, \mathbf{c}^{c2v}), \end{aligned} \quad (6-6)$$

其中  $r$  函数包含一个加权模块用于平衡曝光层面和点击层面的归因结果，以及一个多层感知机用于输出最后的转化率。我们将在下文中进行详细介绍。 $\mathbf{x}_{m_i}$  是最后一个触达点的特征向量，它经过了我们在第6.3.2节描述的嵌入层。 $\mathbf{c}^{i2v}$  是用于表示所有用户行为输入中曝光模式的上下文向量， $\mathbf{c}^{c2v}$  是用于表示模型建模的点击模式的上下文向量。

**通过转化率预估学习归因。**我们使用交叉熵作为转化率预估的损失函数

$$L^v = \sum_{i=1}^n -y_i \log \hat{y}_i - (1 - y_i) \log(1 - \hat{y}_i). \quad (6-7)$$

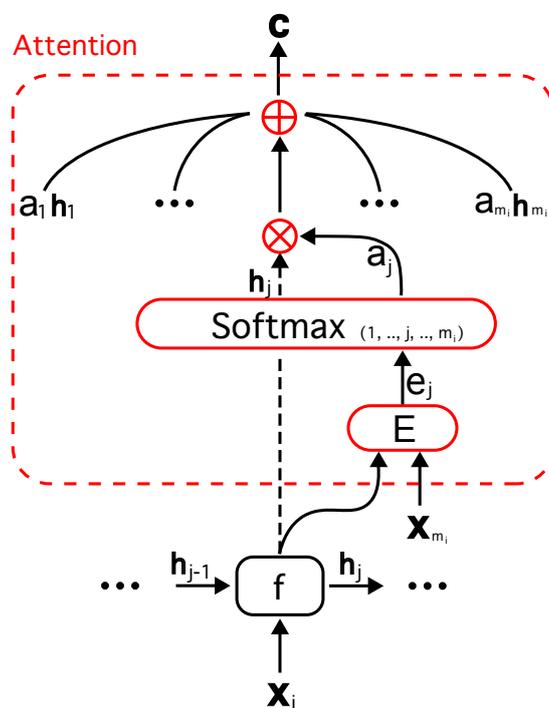


图 6-4 注意力机制的计算过程

Fig 6-4 Attention calculation mechanism

模型的关键组件分别是来自曝光层面和点击层面的注意力输出  $c^{i2v}$  和  $c^{c2v}$ 。图6-4展示了注意力函数  $A$  的结构。

为了计算曝光到转化以及点击到转化的注意力值  $c^{i2v}$  和  $c^{c2v}$ ，我们提出一个统一的基于能量的函数：

$$c = l(\mathbf{h}_1, \dots, \mathbf{h}_j, \dots, \mathbf{h}_{m_i}) = \sum_{j=1}^{m_i} a_j \mathbf{h}_j. \quad (6-8)$$

虽然这个公式是计算曝光层面的隐变量  $c^{i2v}$  的，但不失一般性，我们可以将公式 (6-1) 中的  $\mathbf{x}$  和  $\mathbf{h}$  替换成公式 (6-4) 中的  $\mathbf{s}$ ,  $z$  和  $\mathbf{h}_{m_i}$ ，从而计算点击层面的隐变量  $c^{c2v}$ 。

权值  $a_j$  通过对能量值  $e_j$  做 softmax 操作得到：

$$a_j = \frac{\exp(e_j)}{\sum_{k=1}^{m_i} \exp(e_k)}, \quad (6-9)$$

其中

$$e_j = E(\mathbf{h}_j, \mathbf{x}_{m_i}) \quad (6-10)$$

是一个为每个触达点计算贡献分的能量 (energy) 模型。能量函数  $E$  是一个以  $\tanh$  函数  $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$  作为激活函数的非线性多层神经网络。我们计算注意力值的方法与自然语言处理领域相关文献<sup>[38, 126]</sup> 中的方案类似。

所以，双重注意力机制可以如下表示：

$$\begin{aligned} \mathbf{c}^{i2v} &= \sum_{j=1}^{m_i} a_j^{i2v} \mathbf{h}_j, \\ \mathbf{c}^{c2v} &= \sum_{j=1}^{m_i} a_j^{c2v} \mathbf{s}_j, \end{aligned} \quad (6-11)$$

两式中  $a$  的值可以通过公式 (6-9) 和 (6-10) 计算。

### 基于双注意力机制的归因计算.

我们已经获得了预估的用户转化率  $p(y|\mathbf{x}, \mathbf{z})$  和每个触达点的注意力值  $a_j^{i2v}$  和  $a_j^{c2v}$ 。基于这些计算结果，我们可以很自然地为每一个触达点  $j$  分配归因贡献值。

回顾上文，最终的转化率预估基于模型学习到的双重注意力向量，即  $\mathbf{c}_j^{i2v}$  和  $\mathbf{c}_j^{c2v}$ ，以及最后一个接触点的特征向量  $\mathbf{x}_{m_i}$ 。我们采用一个动态的加权函数  $f_\lambda$  来平衡两个注意力值的贡献

$$\lambda = \frac{\exp[f_\lambda(\mathbf{x}_{m_i}, \mathbf{c}^{c2v})]}{\exp[f_\lambda(\mathbf{x}_{m_i}, \mathbf{c}^{i2v})] + \exp[f_\lambda(\mathbf{x}_{m_i}, \mathbf{c}^{c2v})]}, \quad (6-12)$$

其中  $\lambda$  衡量了点击层面注意力相对于曝光层面注意力的重要性。函数  $f_\lambda$  是一个多层神经网络模型。

所以，我们最初在公式 (6-6) 中提到的预估函数  $r$  可以写成

$$r(\mathbf{x}_{m_i}, \mathbf{c}^{i2v}, \mathbf{c}^{c2v}) = r_{\text{conv}}((1 - \lambda) \cdot \mathbf{c}^{i2v} + \lambda \cdot \mathbf{c}^{c2v}). \quad (6-13)$$

这里  $r_{\text{conv}}(\cdot)$  是一个以 sigmoid 函数为激活函数的多层神经网络。

目前为止，我们已经衡量了曝光层面和点击层面的注意力对最终转化预测的贡献，我们可以很自然地得到每个触达点对最终用户转化的归因贡献值为

$$\text{Attr}_j = (1 - \lambda) \cdot a_j^{i2v} + \lambda \cdot a_j^{c2v}. \quad (6-14)$$

本章使用双重注意力机制的动机是我们希望既曝光层面以及点击层面的用户行为模式来促进转化率预估的准确性，得到更好的归因结果。

### 6.3.4 离线归因评估框架

我们已经得到了用户行为序列中接触点的归因贡献值，我们的焦点转移到如何利用归因结果实现有效的广告预算分配。需要注意的是，除了进行昂贵且含有风险的在线 A/B 测试<sup>[64, 109]</sup> 之外，几乎所有的转化归因相关工作均只报告了其转化率预估的结果，而没有针对归因结果的后续用途做进一步的讨论。在本章我们提出一个离线评估框架，本质上是通过广告活动历史数据对归因模型进行离线评估。

在在线广告中，广告主直观地基于过去的投放表现在广告活动的不同渠道上为后续广告投放分配预算。投放表现在这里指的是投放有效性，即投放到每个频道的广告的投资回报率  $\text{ROI}$ ，该指标通过计算广告活动取得的用户转化和投放成本来衡量。最直接的想法是为  $\text{ROI}$  较高的渠道或子广告活动分配更多预算，以期获得更多的用户转化。但是，不同的归因方法会显著影响  $\text{ROI}$  计算结果<sup>[109]</sup>。具体而言，我们的评估框架的想法是首先计算不同归因模型下每个渠道的  $\text{ROI}$  并根据  $\text{ROI}$  分配不同渠道的预算方案；然后利用广告投放历史记录进行离线重放（replay）模拟，来计算离线预算分配方案的收益，并根据重放时的成本计算不同基线模型在广告投放时的有效性。

接下来我们会先给出基于  $\text{ROI}$  结果的预算分配方法。然后我们会展示我们的重放模拟算法。

**基于  $\text{ROI}$  的预算分配.** 在这个阶段，第一个问题是根据归因结果在  $K$  个广告投放渠道上分配预算  $\{b_1, \dots, b_K\}$ 。这里我们采用相关文献<sup>[109]</sup> 提出的方法：

$$\text{ROI}_{c_k} = \frac{\sum_{\forall y_i=1} \text{Attr}(c_k|y_i) V(y_i)}{\text{渠道 } c_k \text{ 上的总开销}}, \quad (6-15)$$

其中

$$\text{Attr}(c_k|y_i) = \sum_{j=1}^{m_i} \text{Attr}_j \cdot 1(c_j = c_k) \quad (6-16)$$

是聚合所有最终发生转化的用户序列中，投放在渠道  $c_k$  上的广告触达其归因贡献值  $\text{Attr}_j$  的加总。 $1(\cdot)$  是指示函数， $V(y_i)$  是转化能带来的广告主收益常量。计算完  $\text{ROI}$  后我们根据  $\text{ROI}$  的比例在不同渠道上分配预算。对于渠道  $c_k$  来说， $b_k = \frac{\text{ROI}_{c_k}}{\sum_{v=1}^K \text{ROI}_{c_v}} \times B$ 。其中  $B$  是广告主总预算。

**预算重分配的评估算法.** 接下来我们介绍基于预算重分配结果的离线评估算法。所有的广告投放历史数据是一系列事件序列  $\{\text{seq}^s\}$ ，每个序列被表示为  $\{(q_i, \mathbf{x}_i, t_i, y_i, c_i, o_i)\}$ ，其中  $q_i$  是每个用户交互的标识符， $c_i$  是广告触达点所在渠道的编号， $t_i$  是交互时间， $o_i$  是广告话费，特征向量  $\mathbf{x}_i$  包括了点击等标签信息。每个广告触达所导致的用户行为事件要么是一个普通的非转化事件 ( $y_i = 0$ )，要么是一个用户转化事件 ( $y_i = 1$ )。

此外，我们引入了转化黑名单的概念。如果一个渠道  $c_k$  的预算在  $t$  时刻耗尽，那么该用户  $t$  时刻后所有的行为均停止重放模拟，所有广告不再触达；之后所有未完成的广告投放事件序列的转化事件将变为无效，这些转化会被放入转化黑名单，不再计入之后的收益中。这种操作是合理的，因为如果用户无法经历  $t$  时刻之后的广告触达点，则无法保证他最终会在历史用户行为序列重放结束时达成转化。这样的测试结果可以看做是预算重分配方案在真实在线环境中性能表现的下界。

已知  $K$  个渠道上的预算分配结果  $(b_1, b_2, \dots, b_K)$ ，我们可以进行算法6-2所示的重放评估。具体而言，重放过程按照历史时间顺序回顾所有历史广告投放与用户事件。如果频道  $c_k$  的预算在重放进行到事件  $(q_i, \mathbf{x}_i, t_i, y_i, c_i = c_k, o_i)$  时不足以实施该广告展示，则将该序列标识符  $\text{seq}^s$  加入转化黑名单，这个历史用户行为序列的后续所有广告展示与用户行为将不再进行重放。重放完成后，我们可以根据花费  $O$  和取得的用户转化数  $Y$  对（不同转化归因模型得出的）预算重分配结果进行评估。

---

#### 算法 6-2 预算重分配评估算法

---

**Input:** 以历史发生时间  $t_i$  排序的事件序列  $\{(q_i, \mathbf{x}_i, t_i, y_i, c_i, o_i)\}$ ，根据转化归因模型的结果所做的预算分配方案  $\{b_1, \dots, b_K\}$ 。

**Output:** 总转化数量  $Y$  和总开销  $O$ 。

- 1: 初始化转化黑名单  $\mathcal{B} = \{\}$ 、总转化数目  $Y = 0$  总开销  $O = 0$ 。
  - 2: **for** 每一个历史事件  $(q_i, \mathbf{x}_i, t_i, y_i, c_i, o_i)$  **do**
  - 3:     **if**  $\text{seq}^s$  不在转化黑名单  $\mathcal{B}$  之中 **then**
  - 4:         **if** 如果渠道  $c_i$  的剩余预算  $b_{c_i} > o_i$  **then**
  - 5:              $O = O + o_i$
  - 6:              $Y = Y + y_i$
  - 7:              $b_{c_i} = b_{c_i} - o_i$
  - 8:         **else**
  - 9:             将该历史行为序列  $\text{seq}^s$  放入转化黑名单  $\mathcal{B}$  中
  - 10:         **end if**
  - 11:     **end if**
  - 12: **end for**
- 

## 6.4 实验分析

在本小结，我们会先阐述实验的设置，包括数据集的介绍、评价指标和实验中对比的几个模型。然后我们会展示两阶段实验的实验结果。第一个阶段评估模型转化率预估的准确性；第二个阶段我们将评估归因模型指导的预算分配方案的有效性。

### 6.4.1 数据集

我们在两个真实数据集上测试了我们的模型和基线模型。

秒针是一家中国的在线营销技术公司。他们公开的数据集 **Miaozhen**<sup>[113]</sup> 包括从 2013 年 5 月 1 日到 6 月 30 日的大约 12.4 亿广告展示数据。数据中包括大约 5900 万用户和

1044 条用户转化行为。这些广告内容通过 2498 个渠道以 40 种形式（例如搜索引擎、网页、社交媒体等）投放。在这个数据集中，用户每经历一次广告曝光或点击一条广告，其行为的准确时间和附加信息就会被记录。此外，该数据集也包括了用户的购买信息和转化发生的具体时间戳信息。数据集中用户通过匿名的用户标记（cookie）标识符进行跟踪与记录。有了这些记录，我们可以重建用户行为序列，其中包括每个序列中的曝光和点击行为，广告曝光渠道和是否转化的标签信息。

Criteo 是一家研究与开发在线广告相关技术的国际公司，他们在论文<sup>[108]</sup>中公开了 Criteo 数据集以支持实时竞价广告场景下的广告归因分析研究。该数据集由 30 天内的 Criteo 实时流量数据构成，其中包含在 700 多个广告活动中多于 1600 万的广告曝光和 4 万 5 千条转化数据。此数据集中的广告曝光可能会导致用户点击，因此用户行为序列中的每个触达点都会标注是否发生了点击，如果该广告触达导致了转化事件，也会标记转化的相关信息。每一条曝光记录也包括了广告开销信息。这些信息被我们用来进行第二阶段预算分配评估的实验。因为数据集中不存在投放渠道的信息，我们将不同的广告活动看作广告投放与预算分配的不同渠道。

#### 6.4.2 实验流程与评估指标

在此节我们会阐述具体的实验流程和评估指标。我们进行了两个阶段的实验。

第一个阶段着眼于转化率预估性能的评估，这种评估方式在归因任务中有广泛的应用<sup>[113, 114, 117]</sup>。具体而言，给定测试集中的样本  $\{u_i, \{q_j\}_{j=1}^{m_i}, y_i, T_i\}_{i=1}^n$ ，模型预测用户经历一系列触达点序列后，其最终的转化概率是  $\hat{y}_i$ 。我们用两个指标评估模型的表现。**Log-loss** 是事件概率预测的常用指标，计算方法是在前文公式 (6-7) 中给出的交叉熵，数值越低代表性能越好。另一个指标是 **AUC**。该指标衡量模型在分类问题中对一对正负例样本正确排序的能力，即正例的预估转化率应大于负例的预估转化率。**AUC** 数值越高表示性能越好。

第二阶段旨在评估模型归因结果指导多个广告渠道进行预算分配的性能优劣。根据算法 6-2，我们根据时间戳重放所有测试集的广告数据，并计算下面一些指标。注意我们为每个模型设定测试预算为训练集中统计得到的总预算的 1/2, 1/4, 1/8, 1/16, 1/32 倍，这种操作方案在前文第四章中使用过，相似的设定在在线广告相关研究中已经被广泛采用<sup>[9, 16, 23]</sup>。**转化数 (Conversion Number)** 是重放算法输出的所获转化行为的总数，**利润 (Profit)** 是所有转化带来的收益与开销成本之差。另外两个指标是 **CVR (转化率)** 和 **CPA (单次转化成本)**。**CVR** 是用户最终转化的序列占重放时触达用户的广告展示序列数量的比例。这反应了归因模型能为后续广告投放带来的回报率。**CPA** 是平均平均单次转化所花费的成本，衡量了广告活动的效率。因为只有 Criteo 数据集中包含广告花费信

息，所以第二阶段实验只在 Criteo 数据集上进行。

### 6.4.3 对比模型

在这一章节，我们会描述基线模型和各模型的实验设定。我们将本章提出的 DARNN 与其他四个基线模型对比。并讨论了双重注意力机制相比于普通的单独注意力机制的优势。注意，作为我们模型的一个预测输出的监督信号，点击标签  $z$  也被当做输入包含于所有基线模型的输入特征  $\mathbf{x}$  中，这保证了对比实验的公平性。

**LR** 是 Shao 等人<sup>[111]</sup>提出的逻辑斯蒂回归模型，归因贡献值通过训练后模型参数中不同广告投放渠道的系数计算。

**SP** 是 Dalessandro 等人<sup>[112]</sup>提出的一个简单概率模型。我们采用另一文献<sup>[113]</sup>中的方法，每个用户的转化率用如下公式计算：

$$p(y = 1|\{c_j\}_{j=1}^{m_i}) = 1 - \prod_j (1 - \Pr(y = 1|c_j = k)), \quad (6-17)$$

其中  $\Pr(y = 1|c_j = k)$  是在观察到的数据中第  $k$  个渠道上的转化率。

**AH** (AdditiveHazard) 模型是第一篇利用生存分析和加性转化风险函数 (hazard function) 来预测转化率的工作，模型同时考虑了触达点的时间。细节请参考原文献<sup>[113]</sup>。

**AMTA** (Additional Multi-touch Attribution) 模型在<sup>[114]</sup>中被提出，是当前转化归因问题的最优模型。它利用生存分析理论进行转化率预估，并根据转化时刻的风险函数进行转化归因。

**ARNN** 是一个带有一个单独注意力结构的循环神经网络。它仅基于曝光层面的模式建模转化率  $\hat{y}_i = r'(\mathbf{c}^{i2v})$ 。这个模型是为了展示我们的双重注意力结构对于解决稀疏性问题和多角度建模有很好的效果。

**DARNN** 是我们提出的基于双重注意力机制的循环神经网络模型，具体结构在第6.3节已有详细描述。

所有的深度模型都在一台搭载 NVIDIA GTX 1080 Ti 显卡和 Intel Core i7 处理器的计算机上训练，训练时间限制均为五个小时以下。

### 6.4.4 转化率预估实验性能

我们首先对模型在转化率预估任务上的表现进行评估。表6-1 展示了不同模型的具体指标结果。可以看出我们的模型在两个指标上都优于基线模型。这些结果也反映出以下事实：(1) 所有基于注意力机制的模型，即 DARNN 和 ARNN，都取得了较其他模型更强的序列预测结果，这展现了深度神经网络较强的模式识别能力。(2) 基于点过程的

表 6-1 转化率预估实验的评估结果

Fig 6-1 Conversion estimation results on two datasets

Models	Miaozhen		Criteo	
	AUC	Log-loss	AUC	Log-loss
LR	0.8418	0.3496	0.9286	0.3981
SP	0.7739	0.5617	0.6718	0.5535
AH	0.8693	0.6791	0.6791	0.5067
AMTA	0.8357	0.1636	0.8465	0.3897
ARNN	0.8914	0.1610	0.9793	0.1850
DARNN	<b>0.9123</b>	<b>0.1095</b>	<b>0.9799</b>	<b>0.1591</b>

方法如 AH 和 AMTA 在转化率预估任务中表现较差。因为这些模型的主要目标是建模每个触达点处的转化风险率，也并没有考虑用户行为的序列模式信息，从而在转化率预估方面缺乏较好的性能。(3) 对于 log-loss 指标，基线模型获得比深度模型相对更差的结果，这反映出这些基线的预测值与真实值具有较大偏差。相对而言，AUC 不受预测绝对值高低的影响而只衡量成对排序的表现。所以几乎所有基线模型都得到了相对可接受的 AUC 结果。

因为我们提出的 DARNN 捕捉曝光和点击两个层面的模式，并针对两个损失函数，即公式 (6-5) 和 (6-7)，进行优化。所以如图 6-5 所示，在模型的训练过程中产生两条学习曲线。在模型训练过程中，我们首先让模型学习点击信号相关的行为特征模式，即只针对公式 (6-5) 中的点击率预测的损失函数进行优化。当该目标函数收敛，我们开始进行转化率预估的训练，即针对两个损失函数共同训练直到转化率预估的损失函数收敛。当损失函数的值连续上升两次即视为模型收敛与停止训练的条件。之所以采取两阶段训练是使模型在两个训练目标上保持稳定的优化。在推荐系统等领域的多任务学习任务中也有相关文献<sup>[127]</sup> 采用过类似的训练方法。

从图 6-5 中我们可以明显看出本章提出的 DARNN 模型不仅优化了点击预测任务，也优化了转化率预估任务。并且在第二阶段都达到收敛。模型在这两个任务上都取得了很好的效果。

#### 6.4.5 预算重分配离线评估

在实验的第二阶段，我们评估了预算重分配的有效性，这些预算根据不同归因模型输出的归因贡献值来进行重分配，如公式 (6-15) 所示。根据算法 4-1 所示重放历史数据以后，我们计算了各个模型获得的总转化数和总开销花费。为了计算每个模型取得的利润，我们取公式 (6-15) 中每个转化的转化值  $V(y_i = 1)$  为训练集合中统计的平

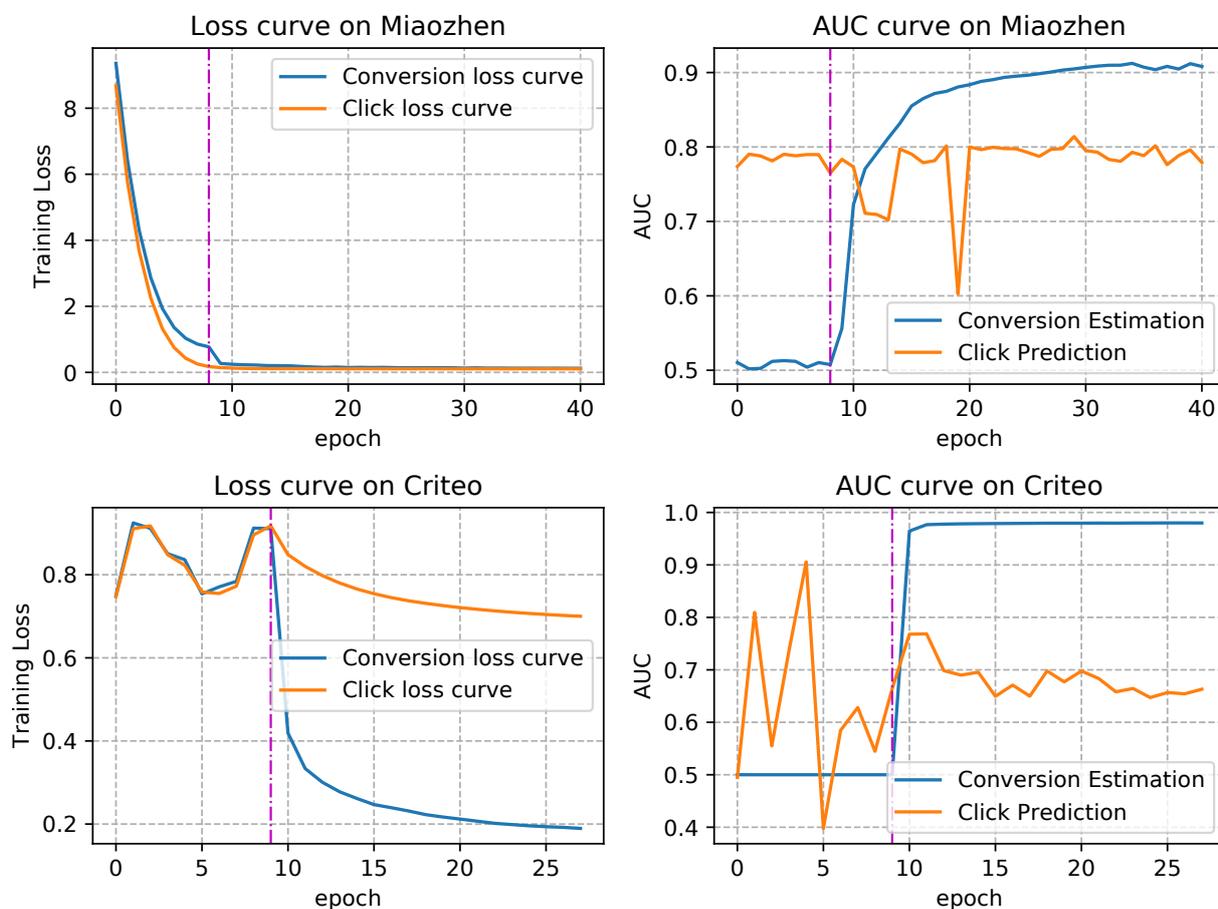


图 6-5 转化归因模型的学习曲线 (点击率预估损失函数值与转化率预估损失函数值)

Fig 6-5 Learning curves of DARNN model on two datasets (CTR loss and CVR loss)

均转化开销这对于所有模型是一个常数，在训练集合数据上的计算方法为  $V(y_i = 1) = (\text{total cost}/\text{conversion number})$ 。第二阶段实验的详细评估结果在表6-2和图6-6中展示。从表格中可以看出，LR 模型的表现最差，所以我们在图6-6中删除了 LR 模型的曲线使得图更易阅读。注意我们在第二阶段的实验中也与基于规则的方法（例如所有转化归因都归于最后一次曝光）进行了对比。我们没有列出其结果，因为 AH 模型与最后一次曝光的方法性能表现几乎相同，我们认为这是一个值得进一步研究的有趣现象。

从表6-2和图6-6中我们可以看出：(1) 当预算增长时，所有模型都会消耗更多成本以获得用户转化，即 CPA 值有所增加。(2) 基于注意力机制的模型，即 ARNN 和 DARNN 都取得了相对其他模型更好的效果。这些提升可能是因为其对于序列模式的挖掘有助于更好的归因。(3) AMTA 和 AH 模型取得了十分相似的结果，这可能是因为他们采用相似的加性转化风险率进行归因建模<sup>[113, 114]</sup>。(4) DARNN 在 CPA 和 CVR 指标

表 6-2 预算重分配的离线评估实验结果

Fig 6-2 Budget allocation evaluation results

模型	CPA					利润 (Profit)				
	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
LR	31.79	29.47	29.77	27.83	27.46	8.022	6.938	4.386	3.238	1.954
SP	24.84	22.98	21.29	21.39	20.93	13.07	10.28	7.694	4.648	2.776
AH	24.69	21.84	20.37	18.89	19.32	27.03	22.08	15.38	10.32	5.491
AMTA	24.71	21.91	20.43	18.89	19.41	27.01	21.96	15.29	10.33	5.446
ARNN	26.66	23.98	22.61	19.86	18.96	29.10	<b>23.32</b>	15.81	11.68	<b>7.010</b>
DARNN	<b>23.47</b>	<b>21.24</b>	<b>18.50</b>	<b>16.85</b>	<b>17.63</b>	<b>29.25</b>	22.56	<b>17.58</b>	<b>12.09</b>	6.26

模型	转化数 (Conversion Num.)					CVR				
	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
LR	576	427	275	181	107	0.0928	0.0910	0.0873	0.0827	0.0748
SP	452	315	191	112	62	0.1205	0.1251	0.1223	0.1122	0.1028
AH	1286	925	607	385	208	0.1120	0.1194	0.1197	0.1183	0.1079
AMTA	1285	922	605	385	207	0.1118	0.1192	0.1195	0.1183	0.1073
ARNN	<b>1527</b>	<b>1073</b>	<b>684</b>	<b>452</b>	<b>262</b>	0.1073	0.1137	0.1119	0.1206	0.1174
DARNN	1315	922	646	419	223	<b>0.1226</b>	<b>0.1274</b>	<b>0.1339</b>	<b>0.1321</b>	<b>0.1206</b>

上取得了最好的结果，这反映了我们提出的双重注意力机制进行归因分析的有效性。同时，实验数据也直观体现了双重注意力机制相对于 ARNN 的单独注意力可以取得更好的效果。(5) ARNN 消耗预算的方式相比其他模型更加激进，所以得到了较差的 CPA 结果。其原因可能在于 ARNN 的归因仅基于广告曝光层面，模型捕捉到的模式倾向于对用户进行长期广告投放，使得成本开销增加。然而，我们的 DARNN 模型使用更加经济的方式进行广告投放，在 CPA 指标上取得了更好的效果，这说明结合曝光层面和点击层面的注意力可以捕捉到长期（曝光到转化）和短期（点击到转化）两种用户行为模式。

#### 6.4.6 可视化分析

在本节，我们将探究我们的归因模型所得出的转化归因结果。首先我们会从触达点角度和渠道角度分别讨论模型得到的归因贡献分的分布情况，然后分析 DARNN 模型根据公式 (6-12) 学习到的用于控制两类用户行为（曝光和点击）影响力的加权参数  $\lambda$ 。

首先，我们展示在 Miaozen 数据集上具有相同长度的序列中平均每个触达点获得的归因贡献分。具体而言，第  $j$  个位置的贡献分的计算方式为在所有最后发生转化并且长度为  $m$  的用户行为序列上的第  $j$  个位置的归因结果取平均  $\overline{\text{Attr}}_j = \frac{1}{N_m} \sum_{i=1}^{N_m} y_i \cdot \text{Attr}_{ij}$ ，其中  $N_m$  是长度为  $m$  的序列的数量， $y_i$  是标记序列是否转化的标签。

图 6-7 展示了长度为 5 和 10 的序列上的归因结果统计。因为 LR 和 SP 模型是直接计算不同渠道的归因值，所以我们无法在图中展示这两个模型在触达点角度的归因结

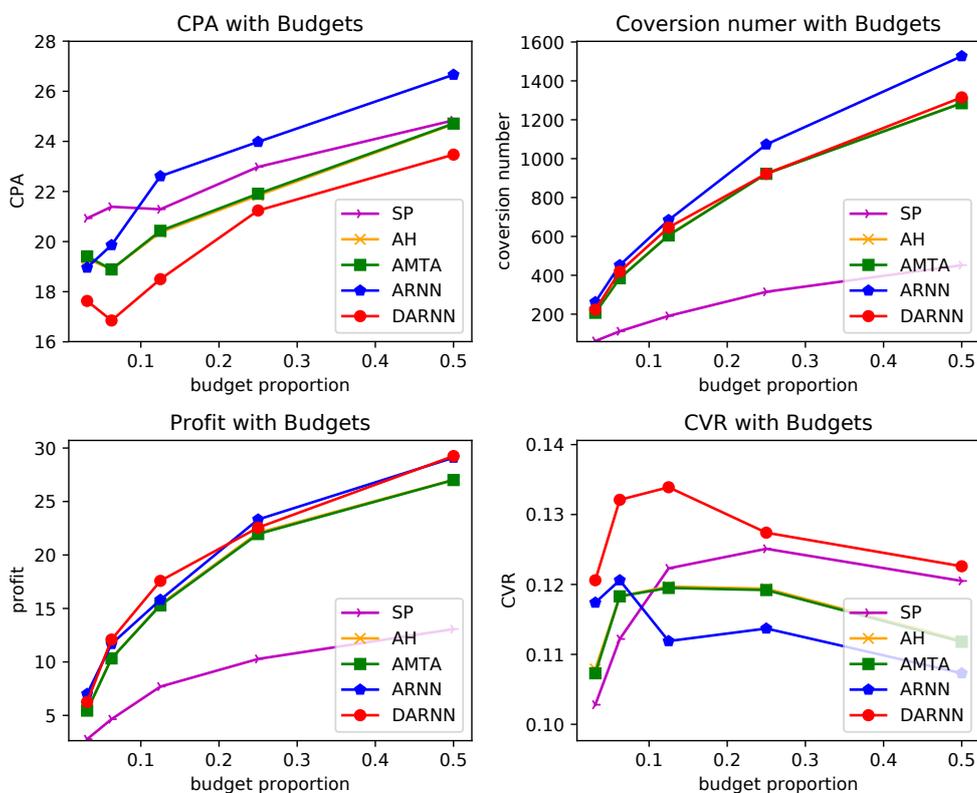


图 6-6 转化归因指导的预算重分配评估实验结果

Fig 6-6 Evaluation performance of budget allocation

果。从图中我们可以看到不同的模型为触达点分配的贡献分存在很大不同。当序列长度较短（长度为 5 的序列）时，DARNN 认为越靠近序列结尾的触达点越可能导致转化。在较长的序列中（长度为 10 的序列），我们的 DARNN 模型为序列中段的触达点分配了更多贡献分，然后在一小段回落后分配了更多的贡献分在靠近最终转化的触达点上。这个现象是合理的，因为无休止的广告投放并不一定是有效的，这反映了本章提出的归因模型在广告投放有效性和用户体验之间的良好权衡。然而 ARNN 似乎给不同触达点平均分配了贡献分。注意到 ARNN 只考虑了曝光层面的归因贡献，这也说明了我们的模型中利用的点击层面的用户行为模式对归因结果有着巨大影响。

我们接下来根据每个渠道获得的归因贡献进行模型分析。在图6-8中我们展示了秒针数据集上发生转化的序列在不同渠道的归因结果。横轴给出了渠道的信息，包括社交媒体和音乐平台等。因为音乐渠道上没有产生过用户转化行为，所以只有 LR 模型在该渠道上分配了归因贡献分，因为该模型将学习到的参数取值来作为不同渠道的归因结果。从图中我们可以看出 (1) LR, SP, AH 和 DARNN 模型都在搜索 (Search) 渠道分配了最高的贡献分，而 ARNN 和 AMTA 在视频 (Video) 渠道分配了最多的贡献分。(2) SP

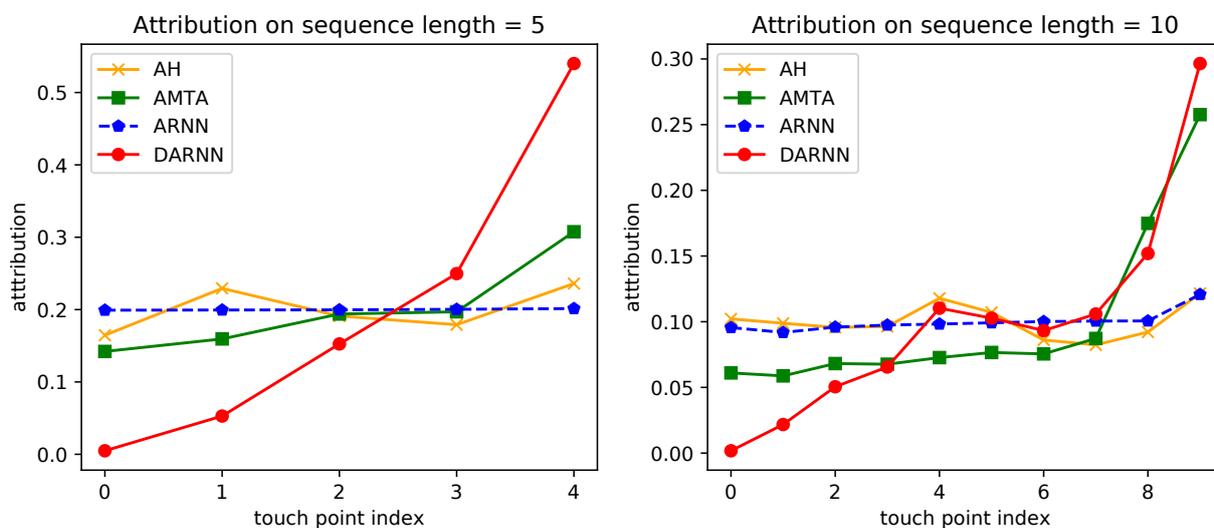


图 6-7 触达点角度的归因统计结果 (Miaozen 数据集)

Fig 6-7 Touch point level attribution statistics (Miaozen)

和 AH 模型在搜索渠道分配的贡献分相比其他模型高很多，而其他模型分配贡献分时则相对平均。(3) 垂直 (Vertical) 和社区 (Community) 渠道被分配到的贡献较少，而基于注意力的模型分配到联合渠道的贡献明显更高。从这些发现我们可以不同模型在给不同广告投放渠道分配贡献分时差异巨大，这也从侧面反映出第二阶段重放实验的必要性。

我们也在图6-9中展示了 Criteo 数据集上的归因结果，并在图6-10中展示了不同模型根据公式 (6-15) 算出的每个渠道 (对于 Criteo 数据集是广告活动) 的 ROI。我们的 DARNN 模型在渠道 5 上分配了最多的贡献分。然而，ROI 计算表明所有模型都在渠道 7 上投入了最高的预算。这个现象的原因可能如公式 (6-15) 所描述的，ROI 不仅与渠道层面归因相关，也和触达点层面的归因有关。

最后，在图6-11中，我们统计了控制曝光层面和点击层面权值的  $\lambda$  的分布。注意，如公式(6-12) 所示，当  $\lambda$  的值变大，点击层面的模式对最后的转化归因会产生更大的影响。我们可以从图中看出点击模式对最终的转化预估影响较大，这也表明了我们双重注意力机制的合理性。

综上所述，实验结果均表明了双重注意力机制结合曝光和点击两种行为模式进行联合建模与预测的重要性，尤其在预算有限情况下，点击层面的行为模式对于最终结果有更高的影响力。

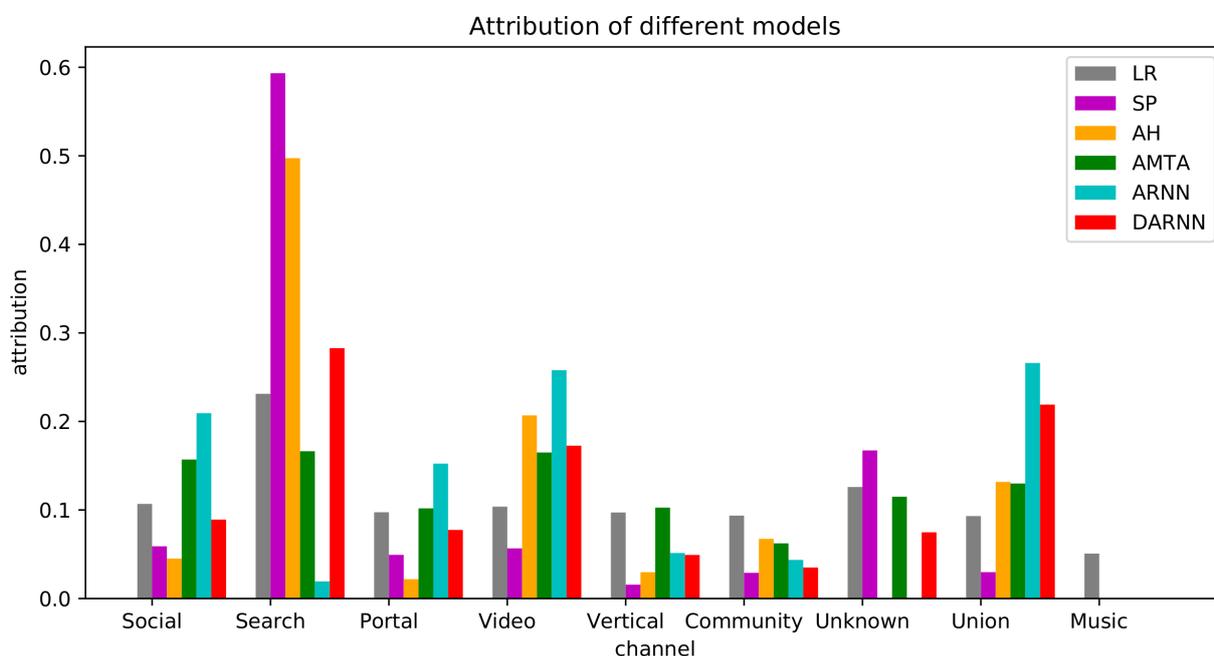


图 6-8 Miaozen 数据集上不同渠道的归因结果

Fig 6-8 Attribution of different channels on Miaozen

## 6.5 本章小结

本章提出了一种基于双重注意力机制的循环神经网络模型，用来在用户行为序列上进行多触达转化归因。我们的模型不仅捕捉到了用户行为的序列模式，也结合了曝光和点击层面的模式进行有效的转化归因。实验结果展示了我们的模型相对当前的已有模型无论在转化率预估还是转化归因合理性方面都有显著提高。本章工作的一个局限性在于我们在计算注意力时没有考虑广告的花费。正如第三章所阐述的预估模型，将成本与市场竞争因素纳入建模中能有效提高成本效益性能。所以在实时竞价广告中，我们计划未来结合竞价策略与市场价格分布等信息进行综合建模与归因。

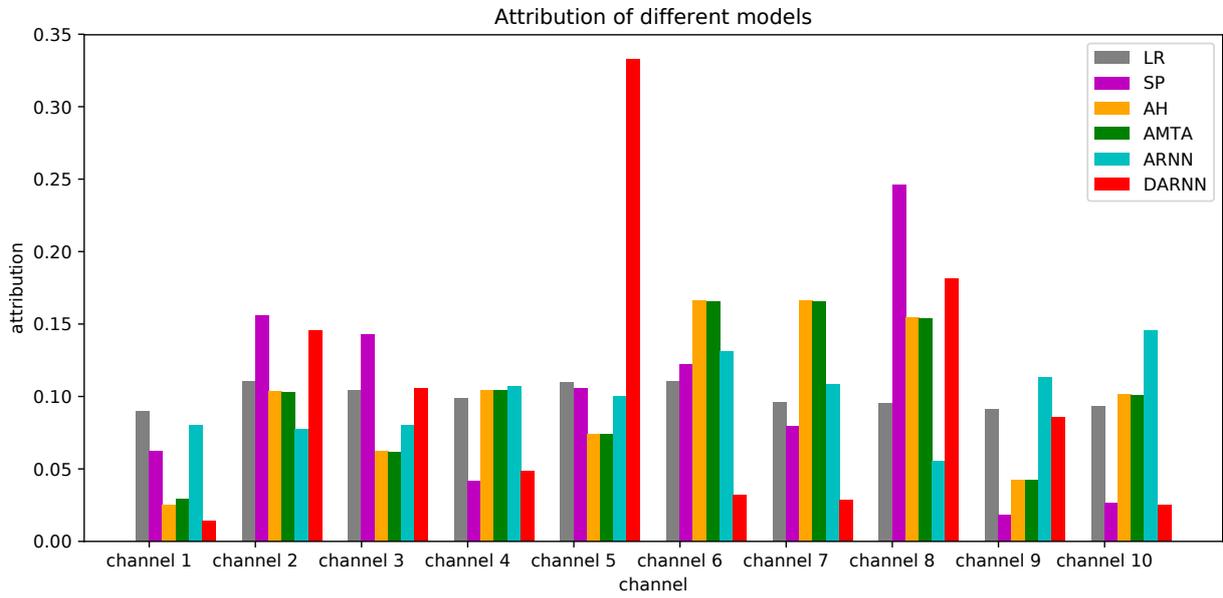


图 6-9 Criteo 数据集上不同广告活动的归因结果  
Fig 6-9 Attribution distribution over campaigns on Criteo

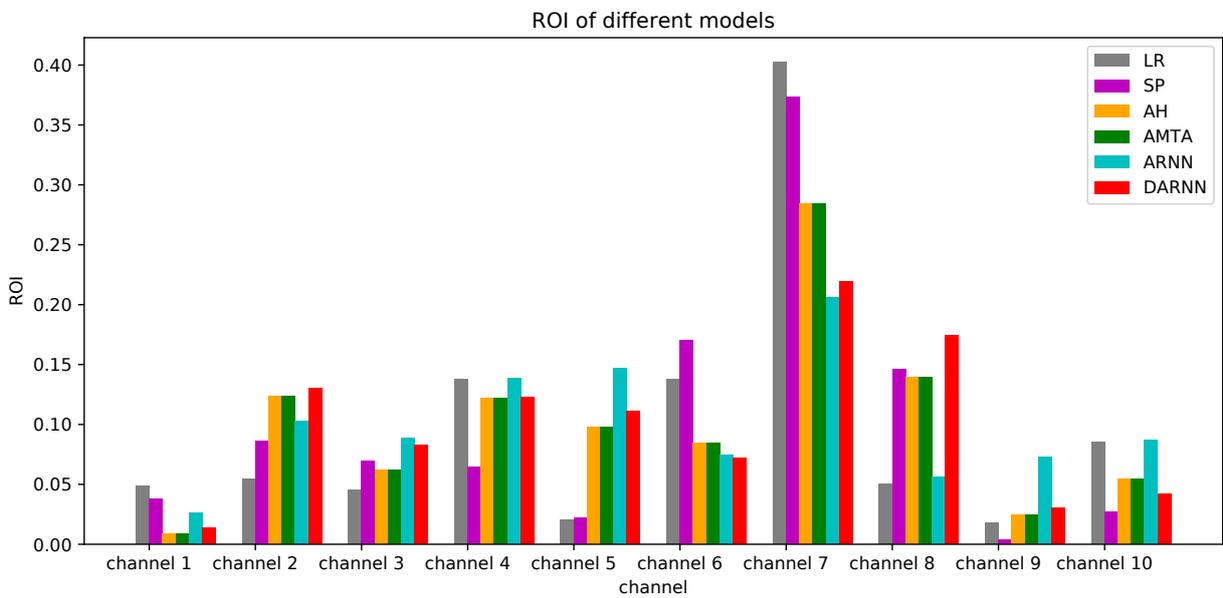


图 6-10 归因模型的归因结果计算所得的 ROI 分布  
Fig 6-10 Calculated ROI over campaigns of Criteo

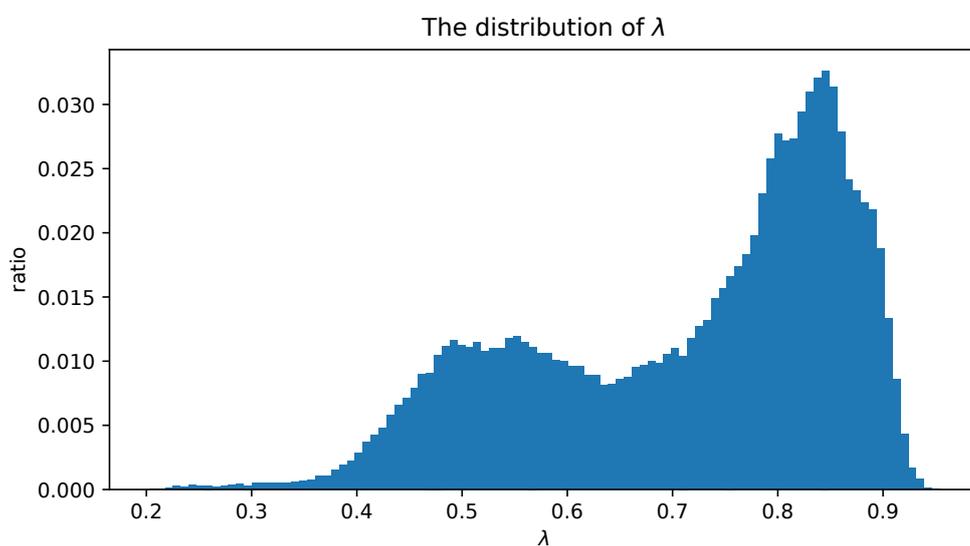


图 6-11 Criteo 数据集上的  $\lambda$  结果  
Fig 6-11 The distribution of  $\lambda$  over Criteo dataset

## 第七章 总结与展望

在线广告作为一种新型的营销形式，在数字媒体与互联网服务行业发挥了重要的作用。其中，实时竞价广告作为一种高效的流量售卖与产品营销的组织形式，正在逐渐受到越来越多的关注，成为在线广告的最受欢迎的计算广告形式。本文针对在线广告领域，特别是实时竞价广告场景，进行了综合性建模并给出了若干解决方案。

我们在第一章详细介绍了在线广告的背景，从数字营销与计算广告入手，阐述了在线广告产业的发展与计算广告的诞生。同时，这一章针对全文的研究内容与研究意义进行了分析说明。本文从实时竞价广告的三方面（用户反馈预测、市场价格预测、竞价策略优化）介绍了实时竞价广告中的建模方法与决策优化机制，并探讨了实时竞价广告中的已知问题与解决办法，为在线广告的建模与优化，及其他智能信息系统提供了研究与建模的新思路。另外，本文还针对另外一些重要的研究课题提出了见解，包括用户序列行为建模、转化归因分析方法与实时竞价广告中的博弈与均衡分析。

第二章主要介绍了实时竞价广告中的市场建模与成本预估任务，从交易机制入手，阐述了实时竞价广告中的竞价形势预估的内容。针对现有工作预设分布假设等存在的问题，同时考虑到二价交易机制的特殊性及其导致的真实数据删失问题，本文提出了基于循环神经网络的新型深度学习方法，用于建模市场价格的分布与估计竞拍胜率。在建模中，我们同时考虑了时序特征模式与条件概率预测，还引入了生存分析方法处理删失问题。本文提出的模型不仅在已有的数据集上超越了现有方法，更是在建模准确性、灵活性等方面具备独到的优势。

第三章针对计算广告中的用户建模与效益预估提出了一种全新的解决方案。以往的方法仅将用户反馈预估看作一种二分类问题并做独立的建模，而并没有考虑预测结果在后面阶段的使用目的及总体优化目标。本文提出将用户反馈建模预测、实时竞价策略的优化置于统一的优化函数下，直接优化广告主利润。我们的方法不仅在多个真实数据集的离线实验取得了巨大的优势，还在广告公司的真实在线平台上获得了超过 20% 的广告主利润提升。

作为实时竞价框架中的最后也是最为重要的一步，竞价策略承担了出价决策的关键作用。在这个步骤中，竞价策略将用户反馈预估即效益预估、竞价形势预估即成本预估的结果作为输入，输出最优的出价决策；同时还需要考虑到广告主预算的限制条件。第四章延续第三章的优化思路，以广告投放收益为优化目标，通过一种统一的优化框架，联合优化效益预估、成本预估与竞价策略。该框架在实际场景中显著提高了广告投

放的效果。在本文附录中，我们还证明了二价竞价机制下的最优竞价策略函数，并分析了博弈均衡。

智能信息系统中针对用户进行建模是非常关键的一个环节。用户建模不仅能够刻画用户行为背后的兴趣与模式，更能为后续用户行为预测提供重要的线索，也就能帮助信息系统的决策提供重要的思路，例如提升广告投放的效果，推进推荐系统的性能改进等。第五章针对线上平台丰富而又多样的用户行为提供一种终生用户建模框架，通过层级的周期记忆网络建模、记忆与存储用户在不同时期、不同场景中的行为模式，为任何时候用户行为预测提供了一种综合建模方案。

另外，考虑到用户在网上浏览过程中不同的广告或商品触达对于用户最终转化购买等行为的具有不同的影响。分析获得广告触达有助于分析广告投放过程的投放效率，更能帮助指导后续广告投放的策略。第六章着眼于用户转化归因问题，针对现有解决方案的问题，提出基于深度学习与注意力机制的转化归因方法，并首次采用一种离线评估框架，用以更高效地指导后续的广告投放活动。

综上所述，本文中对实时竞价广告为主的计算广告研究领域开展了较为全面且具有突破性的研究工作，针对领域中一系列具体的建模与优化问题设计了相应的模型并取得了较为显著的提高，为计算广告领域的进展做出了一定的贡献。本文所涉及的各部分建模方案相辅相成，分别涵盖预测、决策两大应用方向，解决方案基本涵盖了机器学习领域的几个典型问题：分类、回归和优化。

此外，计算广告的研究方法也适用于很多其他领域，包括推荐系统、搜索引擎等智能信息系统。其本质区别在于前者更多时候是带约束的优化问题，而其他方向的任务并没有显式的约束，以至于无需考虑很多限制条件进行优化。

然而，计算广告领域尚存在众多研究课题等待定义与解决，而本文提出的建模与解决方案亦存在可待改进的空间。未来工作将着眼于以下几方面：

- 结合第二章所提出的生存分析模型与第四章讨论的联合优化框架，更好地处理后者模型结构中未曾解决的删失数据问题；
- 将第六章的优化框架构建于广告投放收益的优化思路之上，重新思考转化归因模型对于广告投放收益的影响<sup>[108]</sup>，而不仅仅从转化率预估的角度进行归因分析；
- 将第五章的模型用于其他智能信息系统的任务中，例如推荐、搜索等，并引入任务相关的优化目标函数，提高本文提出的用户建模方法的通用性。

总而言之，实时竞价广告中的建模与决策优化作为机器学习等科学方法应用于工业场景最为前沿的领域，其包含的建模思想与优化方法为所在行业及相关领域的任务提供了较多的启发，同时还有较多的课题等待计算机科学、数学、经济学等领域的专家进行持续的探索与发现。

## 附录 A 竞价机器算法的在线更新公式

第四章介绍了竞价机器优化算法。模型参数的更新有多种方案，例如随机梯度下降法 SGD<sup>[19]</sup>、前后分割更新算法<sup>[128]</sup> (Forward-Backward Splitting, FOBOS) 以及归整双平均算法<sup>[129]</sup> (Regularized Dual Averaging, RDA)。为了在该算法中实现在线学习的模型参数更新范式，我们推导了 FTRL<sup>[51]</sup> (Follow-the-regularized-Leader) 算法，使得模型能够动态控制学习过程中的模型稀疏性，更好地提升泛化性能。

我们使用  $\mathbf{g}_t$  来表示第  $t$  个样本。更新用户反馈预估模型和市场形势预估模型参数的方程是

$$\mathbf{g}_t^\theta = \frac{\partial U_t}{\partial \theta}, \quad \mathbf{g}_t^\phi = \frac{\partial U_t}{\partial \phi}. \quad (\text{A-1})$$

并且  $\mathbf{g}_{1:t} = \sum_{s=1}^t \mathbf{g}_s$ .

在在线梯度下降 SGD 的形式中，两个模型的参数  $\theta$  和  $\phi$  将分别以如下公式进行更新

$$\begin{aligned} \theta_{t+1} &= \theta_t - \eta_t^\theta \mathbf{g}_t^\theta, \\ \phi_{t+1} &= \phi_t - \eta_t^\phi \mathbf{g}_t^\phi, \end{aligned} \quad (\text{A-2})$$

其中  $\eta_t$  是递减的学习率。

我们使用 FTRL 的算法来对参数进行如下更新：

$$\begin{aligned} \theta_{t+1} &= \arg \min_{\theta} (\mathbf{g}_{1:t}^\theta \cdot \theta + \frac{1}{2} \sum_{s=1}^t \delta_s \|\theta - \theta_s\|_2^2 + \lambda_1 \|\theta\|_1), \\ \phi_{t+1} &= \arg \min_{\phi} (\mathbf{g}_{1:t}^\phi \cdot \phi + \frac{1}{2} \sum_{s=1}^t \delta_s \|\phi - \phi_s\|_2^2 + \lambda_1 \|\phi\|_1), \end{aligned} \quad (\text{A-3})$$

这里我们定义学习率的更新方式为  $\delta_{1:t} = \frac{1}{\eta_t}$ 。

正如 FTRL 原文<sup>[51]</sup> 所属，所有系数只有一个取值会被记录，所以我们可以将上述两种参数的  $\arg \min$  方程重写为

$$\begin{aligned} \frac{1}{\eta_t} \|\theta\|_2^2 + (\mathbf{g}_{1:t}^\theta - \sum_{s=1}^t \delta_s \theta_s) \cdot \theta + \lambda_1 \|\theta\| + (\text{常量}), \\ \frac{1}{\eta_t} \|\phi\|_2^2 + (\mathbf{g}_{1:t}^\phi - \sum_{s=1}^t \delta_s \phi_s) \cdot \phi + \lambda_1 \|\phi\| + (\text{常量}). \end{aligned} \quad (\text{A-4})$$

如果我们存储  $\mathbf{z}_{t-1}^\theta = \mathbf{g}_{1:t-1} - \sum_{s=1}^{t-1} \delta_s \boldsymbol{\theta}_s$  (对于  $\phi$  也类似操作), 我们可以针对第  $t$  个样本进行如下更新

$$\begin{aligned} \mathbf{z}_t^\theta &= \mathbf{z}_{t-1}^\theta + \mathbf{g}_t^\theta + \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right) \boldsymbol{\theta}_t, \\ \mathbf{z}_t^\phi &= \mathbf{z}_{t-1}^\phi + \mathbf{g}_t^\phi + \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right) \boldsymbol{\phi}_t. \end{aligned} \quad (\text{A-5})$$

So that we can update per-coordinate of the parameter as

$$\begin{aligned} \boldsymbol{\theta}_{t+1,i} &= \begin{cases} 0 & \text{if } |z_{t,i}^\theta| \leq \lambda_1, \\ -\eta_t^\theta (z_{t,i}^\theta - \text{sgn}(z_{t,i}^\theta) \lambda_1) & \text{否则,} \end{cases} \\ \boldsymbol{\phi}_{t+1,i} &= \begin{cases} 0 & \text{if } |z_{t,i}^\phi| \leq \lambda_1, \\ -\eta_t^\phi (z_{t,i}^\phi - \text{sgn}(z_{t,i}^\phi) \lambda_1) & \text{否则,} \end{cases} \end{aligned} \quad (\text{A-6})$$

这里的第二个下标  $i$  是参数的不同维度坐标的下标。

对于学习率  $\eta_t$ , 我们采用了如下更新形式:

$$\eta_{t,i} = \frac{\alpha}{\beta + \sqrt{\sum_{s=1}^t \mathbf{g}_{s,i}^2}}. \quad (\text{A-7})$$

## 附录 B 最优竞价策略理论证明

在本章，我们基于二价竞价机制，证明最大化广告投放收益的最优竞价函数的具体形态。

**定理 B.1.** 二价竞价机制下，最大化广告投放收益的最优竞价函数是基于预估效益的线性函数。

**证明.** 首先，如第2.1所述，在二价竞价机制下，胜率函数与市场价格分布有如下关系：

$$W(b) = \int_0^b p_z(z) dz. \quad (\text{B-1})$$

$c(b)$  是竞拍价格  $b$  时的期望成本

$$c(b) = \frac{\int_0^b z p_z(z) dz}{\int_0^b p_z(z) dz}. \quad (\text{B-2})$$

我们依旧使用  $r$  来表示用户反馈预估模型的输出结果，那么  $b(r)$  就是竞价策略函数， $u(r)$  是广告主设定的效益函数（例如单次点击期望收益  $u(r) = v \times r$ ）。

我们的优化问题是在广告主预算  $B$  的限制下，最大化利润收益，即

$$\begin{aligned} \max_{b(\cdot)} \quad & T \int_r [u(r) - c(b(r))] W(b(r)) p_r(r) dr, \\ \text{s.t.} \quad & T \int_r c(b(r)) W(b(r)) p_r(r) dr = B, \end{aligned} \quad (\text{B-3})$$

这里的  $T$  是竞价请求的总数量。

那么该优化问题 (B-3) 的拉格朗日函数是

$$\begin{aligned} \mathcal{L}(b(r), \lambda) = & \int_r [u(r) - c(b(r))] W(b(r)) p_r(r) dr \\ & - \lambda \cdot \int_r c(b(r)) W(b(r)) p_r(r) dr + \frac{\lambda B}{T}, \end{aligned} \quad (\text{B-4})$$

$\lambda$  是拉格朗日乘子系数。

**求解**  $b(\cdot)$ 。根据变分法，竞价函数  $b(r)$  的欧拉-拉格朗日条件是

$$\frac{\partial \mathcal{L}(b(r), \lambda)}{\partial b(r)} = 0, \quad (\text{B-5})$$

这可以推导出

$$\begin{aligned}
0 &= \frac{\partial \mathcal{L}(b(r), \lambda)}{\partial b(r)} \\
\Rightarrow 0 &= u(r)p_r(r) \frac{\partial W(b(r))}{\partial b(r)} - (\lambda + 1)p_r(r) \\
&\quad \cdot \left[ \frac{\partial c(b(r))}{\partial b(r)} W(b(r)) + c(b(r)) \frac{\partial W(b(r))}{\partial b(r)} \right] \\
\Rightarrow (\lambda + 1)p_r(r) \frac{\partial c(b(r))}{\partial b(r)} W(b(r)) \\
&= [u(r) - (\lambda + 1)c(b(r))] p_r(r) \frac{\partial W(b(r))}{\partial b(r)}.
\end{aligned} \tag{B-6}$$

因为

$$\frac{\partial c(b(r))}{\partial b(r)} = \frac{p_z(b(r)) \left[ b(r) \int_0^b p_z(z) dz - \int_0^b z p_z(z) dz \right]}{(W(b(r)))^2}, \tag{B-7}$$

并且

$$\frac{\partial W(b(r))}{\partial b(r)} = p_z(b(r)), \tag{B-8}$$

将公式 (B-7) 和 (B-8) 代入公式 (B-6)，我们可以推导出欧拉-拉格朗日条件变成了

$$\begin{aligned}
(\lambda + 1)b(r) &= u(r) \\
\Rightarrow b(r) &= \frac{u(r)}{\lambda + 1}.
\end{aligned} \tag{B-9}$$

□

## 附录 C 最优竞价策略的博弈分析

本章将从博弈论的角度，详细阐述市场上多个广告主进行竞价博弈时，最优竞价策略与对应的博弈均衡理论。

### C.1 基于收益优化的博弈分析

在本节，我们根据相关文献<sup>[31, 130]</sup>中对于实时竞价广告的一价竞价模式下最优化用户点击数目的证明，研究带预算约束的对称博弈在实时竞价广告中的理论。首先，我们推导了实时竞价广告中考虑广告主收益的优化目标下，博弈均衡对应的最优竞价策略函数。然后，基于推导得到的竞价函数，我们讨论了“公地悲剧” (tragedy of the commons) 陷阱在实时竞价广告场景中的体现。虽然相关文献<sup>[31, 130]</sup>已经针对一些类似情况进行了分析讨论，但是在本章，我们讨论的内容与这些已有工作略有不同，包括竞价机制、优化目标等方面的区别。我们讨论的内容主要是关于二价竞价机制下最优化广告主收益的目的。

#### C.1.1 问题设定

我们首先讨论一些符号与概念的设定。我们首先为所有的概率密度函数与累积分布函数添加了  $b, z, r$  这些下标以标志是关于这些变量的分布函数。

**竞价函数的单调性.** 在一般的博弈论分析设定<sup>[131]</sup>中，有  $n(n \geq 2)$  个采用相同竞价策略  $b(r)$  的广告主，这些竞价策略以预估的用户反馈  $r$  为输入并输出竞拍价格  $b$ 。对于竞价策略函数  $b(r)$  的输出一般可以认为随着用户反馈概率  $r$  单调递增，即

$$b(r_1) > b(r_2) \Leftrightarrow r_1 > r_2. \quad (\text{C-1})$$

后面我们会证明其单调性。

每次进行广告展示竞拍时，对于每一位广告主的用户反馈概率  $r$  服从  $p_r(r)$  的分布并且是独立同分布 (i.i.d.) 的，其对应的累积分布函数是  $F_r(r)$ ：

$$F_r(r) = \int_0^r p_r(t) dt, \quad \frac{\partial F_r(r)}{\partial r} = p_r(r). \quad (\text{C-2})$$

我们还定义了  $F_b(b)$  作为竞拍价格  $b$  的累积分布函数，也就是竞拍价格小于  $b$  的概率：

$$F_b(b) = \int_0^b p_b(a) da. \quad (\text{C-3})$$

需要注意的是

$$F_b(b(r)) = P(b(r) > b(r_2)) = P(r > r_2) = F_r(r), \quad (\text{C-4})$$

因为  $b(r)$  随着  $r$  单调递增。因此，对于市场价格变量  $z$ ，即任一广告主的  $(n-1)$  竞争对手出价的最高价格，其累积分布函数是  $F_z(z)$ ：

$$F_z(z) = F_b(z)^{n-1}, \quad (\text{C-5})$$

那么它对应的概率密度函数  $p_z(z)$  是

$$p_z(z) = \frac{\partial F_z(z)}{\partial z} = (n-1)F_b(z)^{n-2}p_b(z). \quad (\text{C-6})$$

**对称博弈中的胜率函数.** 对于 1 号广告主的胜率函数  $W_r(r)$ ，不失一般性，其对应的反馈概率  $r$  是所有  $n$  个广告主中最高的：

$$W_r(r) = P(r > r_2, r > r_3, \dots, r > r_n) = F_r(r)^{n-1}. \quad (\text{C-7})$$

根据公式 (C-4) 和 (C-5)，我们可以推导出胜率函数  $W_b(b(r))$ ，和我们正文中推导的胜率函数  $W(b(r))$  含义一致

$$\begin{aligned} W_b(b(r)) &= P(b(r) > b_2, b(r) > b_3, \dots, b(r) > b_n) \\ &= F_b(b(r))^{n-1} = F_r(r)^{n-1} \\ &= P(r > r_2, r > r_3, \dots, r > r_n) \\ &= W_r(r). \end{aligned} \quad (\text{C-8})$$

**期望效益与期望成本.** 我们根据之前公式 (B-2) 的结果然后导出期望利润  $R(r, b)$  如下：

$$R(r, b) = u(r) - c(b) = u(r) - \frac{\int_0^b zp_z(z)dz}{\int_0^b p_z(z)dz}. \quad (\text{C-9})$$

这里的  $b$  是竞价策略函数  $b(r)$  的输出。

### C.1.2 对称博弈的最优竞价策略函数

在已经定义了问题及效益、成本函数之后，我们接下来推导实时竞价广告场景下，优化广告投放收益的约束优化问题的最优解析解。

**定理 C.1.** 在考虑预算限制的重复对称博弈中，最优竞价策略函数与预估的效益有关系，并且其策略输出的竞拍价格关于市场上参与竞拍的广告主数量成单调递增关系。

**证明.** 我们的优化目标是最大化每一位参与竞拍的广告主的利润，同时又需要考虑二价竞价机制下广告主的预算限制  $B$ 。所以问题可以形式化定义为

$$\begin{aligned} \max_{b(\cdot)} \quad & T \int_r [u(r) - c(b(\tau))] W_b(b(\tau)) p_r(r) dr, \\ \text{s.t.} \quad & T \int_r c(b(\tau)) W_b(b(\tau)) p_r(r) dr = B, \end{aligned} \quad (\text{C-10})$$

这里我们假设竞价过程是基于信号  $\tau$  的，而  $\tau$  与预估的用户反馈概率有关。

接下来可以写出拉格朗日函数  $\mathcal{L}(\tau, \lambda)$  如下：

$$\begin{aligned} \mathcal{L}(\tau, \lambda) = & \frac{\lambda B}{T} + \int_r u(r) W_b(b(\tau)) p_r(r) dr \\ & - (\lambda + 1) \int_r c(b(\tau)) W_b(b(\tau)) p_r(r) dr, \end{aligned} \quad (\text{C-11})$$

这里的  $\lambda$  是拉格朗日乘子系数。需要注意的是，效益函数  $u(r)$  只受真实的反馈概率  $r$  的影响，而成本和竞拍价格有关，后者又是基于已知的用户反馈概率  $\tau$ 。

考虑公式 (C-8) 和 (C-9) 拉格朗日函数可以推导出

$$\begin{aligned} \mathcal{L}(\tau, \lambda) = & \frac{\lambda B}{T} + \int_r \left[ u(r) F_r(\tau)^{n-1} \right. \\ & \left. - (\lambda + 1) \int_0^{b(\tau)} z p_z(z) dz \right] p_r(r) dr, \end{aligned} \quad (\text{C-12})$$

**求解  $b(\cdot)$ .** 我们可以计算目标函数关于  $\tau$  的梯度为

$$\begin{aligned} \frac{\partial \mathcal{L}(\tau, \lambda)}{\partial \tau} = & \int_r \left[ u(r) (n-1) F_r(\tau)^{n-2} p_r(\tau) \right. \\ & \left. - (\lambda + 1) b(\tau) p_z(b(\tau)) \frac{\partial b(\tau)}{\partial \tau} \right] p_r(r) dr. \end{aligned} \quad (\text{C-13})$$

这是一个对称均衡，当自变量取到真值时，即  $\tau = r$ ，优化目标达到最优<sup>[131]</sup>。因此可以得出

$$\begin{aligned} \frac{\partial \mathcal{L}(\tau, \lambda)}{\partial \tau} = 0 \Big|_{\tau=r} \Rightarrow \\ u(r) (n-1) F_r(r)^{n-2} p_r(r) = (\lambda + 1) b(r) p_z(b(r)) \frac{\partial b(r)}{\partial r}. \end{aligned} \quad (\text{C-14})$$

因为  $b(r)$  关于  $r$  是单调递增的，所以他们各自的概率密度函数  $p_r(r)$  和  $p_b(b)$  有如下性质：

$$p_r(r) = p_b(b(r)) \frac{\partial b(r)}{\partial r}. \quad (\text{C-15})$$

将公式 (C-6) 和 (C-15) 代入公式 (C-14) 得

$$\begin{aligned}
 & u(r)(n-1)F_r(r)^{n-2}p_r(r) \\
 &= (\lambda+1)b(r)(n-1)F_b(b(r))^{n-2}p_b(b(r))\frac{\partial b(r)}{\partial r} \\
 \Rightarrow & u(r)(n-1)F_r(r)^{n-2}p_r(r) \\
 &= (\lambda+1)b(r)(n-1)F_b(b(r))^{n-2}p_r(r) \\
 \Rightarrow & b(r) = \frac{u(r)}{\lambda+1}.
 \end{aligned} \tag{C-16}$$

我们可以发现，推导得出的最优竞价策略函数与效益函数  $u(r)$  成线性关系。

特别地，在正文第四章，我们采用了如公式 (4-2) 所示的效益函数

$$u(r) = vr, \tag{C-17}$$

这里的  $v$  是单次用户正反馈的正收益。因此我们可以得到

$$b(r) = \frac{vr}{\lambda+1}. \tag{C-18}$$

到目前为止，我们已经推导出了重复交易的对称博弈中最优竞价策略函数。其中所有广告主都采用一样的竞价函数。对比附录B推导得到的单一广告主的最优竞价策略函数，不同之处在于  $\lambda$  的取值。接下来我们将讨论  $\lambda$  的求解。

**求解  $\lambda$ .** 将拉格朗日函数关于  $\lambda$  的梯度代入我们可以得到预算限制方程

$$\begin{aligned}
 \frac{\partial \mathcal{L}(\tau, \lambda)}{\lambda} &= 0 \\
 \frac{B}{T} &= \int_r \int_0^{\frac{vr}{\lambda+1}} z p_z(z) dz p_r(r) dr \\
 \frac{B}{T} &= \int_r \int_0^{\frac{vr}{\lambda+1}} z(n-1)F_b(z)^{n-2}p_b(z) dz p_r(r) dr \\
 \frac{B}{T} &= \int_r \int_0^{\frac{vr}{\lambda+1}} z(n-1)F_r\left(\frac{(\lambda+1)z}{v}\right)^{n-2} p_r\left(\frac{(\lambda+1)z}{v}\right) \\
 &\quad \cdot \frac{(\lambda+1)}{v} dz p_r(r) dr \\
 \frac{B}{T} &= \int_r \int_0^r \frac{vt}{\lambda+1} (n-1)F_r(t)^{n-2} p_r(t) dt p_r(r) dr \\
 \frac{v}{\lambda+1} &= \frac{B}{T \int_r \int_0^r t(n-1)F_r(t)^{n-2} p_r(t) dt p_r(r) dr}.
 \end{aligned} \tag{C-19}$$

第三个等式可以用公式 (C-6) 推导得出, 第四个等式需要考虑公式 (C-4)。因此我们可以获得最新的最优竞价策略函数为

$$b(r) = \frac{vr}{\lambda + 1} = \frac{Br}{T \int_r \int_0^r t(n-1)F_r(t)^{n-2} p_r(t) dt p_r(r) dr}. \quad (\text{C-20})$$

从公式 (C-20) 我们可以发现, 分母是正数并且他关于  $n \geq 2$  的导数为负。这也就表明最优竞拍价格  $b(r)$  是随着参与竞拍的广告主数量  $n$  单调递增。□

**基于特例的解析解.** 我们利用  $p_r(r)$  的一个特例来给出最优竞价函数  $b(r)$  的一个解析解。我们假设用户反馈预估结果  $r$  是服从均一分布, 也就是说

$$p_r(r) = 1, F_r(r) = r. \quad (\text{C-21})$$

因此, 最优竞价策略函数的闭型是

$$\begin{aligned} b(r) &= \frac{vr}{\lambda + 1} = \frac{Br}{T \int_r \int_0^{r'} t(n-1)t^{n-2}(t) dt dr'} \\ &= \frac{Br}{T \int_r \frac{n-1}{n} r'^n dr'} = \frac{Br \cdot n(n+1)}{T(n-1)}. \end{aligned} \quad (\text{C-22})$$

在这种情况下, 我们发现最优竞价函数和每笔交易的平均预算  $\frac{B}{T}$  以及市场上参与竞价的广告主数量  $n$  是线性关系。

### C.1.3 “公地悲剧”

在本章节我们将基于以上推导得出的结果做一些延伸讨论。

我们首先定义性能对比的框架。对于广告主来说, 他会对比广告展示所获得的效益  $u(r)$ , 例如  $u(r) = \sum vy$ ,  $y$  是用户正反馈的标记符。对他来说, 越高的效益越好, 也就是他希望获得越多的用户正反馈。另外, 如果效益相同, 那么所投入的成本越低越好。

从公式 (C-20) 和 (C-22) 中我们可以发现, 最优竞拍价格关于市场上的竞争者数量成单调递增关系, 也就是说任何一位广告主都试图最大化效益并保持总成本小于其额定预算。参与竞拍的广告主数量是  $n$ , 每一个竞拍者都将以  $\frac{1}{n}$  的概率赢得一部分交易。这种情况下, 每一位竞拍者都会倾向于花掉所有的预算来获取最大的目标效益。然而这种均衡是无效的, 因为这将导致社会福利很低下的情形, 也就是说所有参与竞拍的广告主都会穷尽所有预算来赢取不变的效益,  $\frac{1}{n}$  倍的总广告展示机会与总用户正反馈。

一个相对好一点的情形是每一位广告主都耗费  $B/n$  的预算并得到这些不变的效益 (与上述广告展示机会和用户正反馈数目一致)。极端情况下当  $n \rightarrow \infty$ , 每一位广告主

以 0 元的价格竞拍，胜者会随机挑选得出，然后也只需要支付 0 元。但是这种情况是不切实际的，因为每一位竞拍者会和其他人互相竞争而非互相合作。在这种不稳定的状况下，每一位广告主会倾向于出高价来赢得交易，并赢得广告展示机会，以求最大化各自的期望效益，即  $F_z(z)$ 。最后，整个系统会进入一种如公式 (C-20) 和 (C-22) 所示的均衡状态，即每一位广告主都会最终耗尽他的总预算。

这就是“公地悲剧”<sup>[132]</sup> (tragedy of the commons) 在实时竞价广告场景下的一种体现。“公地悲剧”是一种涉及个人利益与公共利益对资源分配有所冲突的社会陷阱。实时竞价广告的公地悲剧反映出带有预算限制的实时竞价环境中的一种“囚徒困境”<sup>[133]</sup>。需要注意的是，公地悲剧不仅存在于我们推导出的最优竞价策略，其实在所有与期望效益成单调递增关系的竞价策略中都存在。更多的讨论可以参考相关文献<sup>[130, 131]</sup>。

## 附录 D 循环神经网络的实现

第二章介绍了基于循环神经网络的深度形势预估模型 DLF。在本章我们将介绍 DLF 的具体神经网络实现细节。

每一条样本是一条三元组  $(\mathbf{x}, z, b_l)$ ，其中  $\mathbf{x} \in \mathbb{R}^K$  是  $K$  维特征向量， $z \in \mathbb{N}^+$  是整数代表真实的市场价格， $b_l \in \mathbb{N}^+$  是也是整数代表竞拍价格。

如图2-5所示，DLF 模型的输入是  $(\mathbf{x}, z, b_l)$ 。对于每一个循环神经网络单元，我们输入  $(\mathbf{x}, b_j)$ ，其中  $1 \leq j \leq l$  是不同循环神经网络单元对应的价格区间下标。

在实时竞价广告中， $\mathbf{x}$  一般是多热编码 (multi-hot encoded) 向量，包括一系列特征信息，所以我们首先将特征向量通过一层嵌入 (embedding) 层

$$\mathbf{e} = \text{embed}(\mathbf{x})$$

，获得嵌入向量  $\mathbf{e}$ ，然后我们将其与价格变量  $b_j$  拼接得到

$$\mathbf{v}_j = \text{concat}(\mathbf{e}, b_j)$$

。对于循环神经网络单元的实现，我们使用了长短期记忆 (Long Short-Term Memory, LSTM) 模型<sup>[37]</sup>，每个单元具体计算过程如下

$$\begin{aligned} \mathbf{f}_j &= \sigma(\mathbf{W}_f \cdot \mathbf{v}_j + \mathbf{b}_f) \\ \mathbf{i}_j &= \sigma(\mathbf{W}_i \cdot \mathbf{v}_j + \mathbf{b}_i) \\ \mathbf{o}_j &= \sigma(\mathbf{W}_o \cdot \mathbf{v}_j + \mathbf{b}_o) \\ \mathbf{r}_j &= \mathbf{f}_j \odot \mathbf{r}_{j-1} + \mathbf{i}_j \odot \tanh(\mathbf{W}_s \cdot \mathbf{v}_j + \mathbf{b}_s) \\ \mathbf{l}_j &= \mathbf{o}_j \odot \tanh(\mathbf{r}_j). \end{aligned} \tag{D-1}$$

这里的  $\mathbf{r}_j$  是第  $j$  个循环单元的隐状态。

经过上述计算过程，我们获得了每个神经网络单元的输出  $\mathbf{l}_j$ ，然后我们将其通过一层全连接层，激活函数是 sigmoid 函数  $\sigma(x) = \frac{1}{1+e^{-x}}$ ，最后输出每个时间点的条件胜率变量  $h_j$  的值

$$h_j = \sigma(\mathbf{W}_h \cdot \mathbf{l}_j + \mathbf{b}_h)$$

，如公式 (2-7) 所述。

最后，我们通过公式 (2-8) 和 (2-9) 计算市场价格分布  $p_z(z)$ 、胜率函数  $W(b)$  以及输率函数  $S(b)$  的取值。

## 参考文献

- [1] GRAEPEL T, CANDELA J Q, BORCHERT T, et al. Web-scale bayesian click-through rate prediction for sponsored search advertising in microsoft's bing search engine[C]// . 2010.
- [2] JANSEN B J, SPINK A. Sponsored search: is money a motivator for providing relevant results?[J]. Computer, 2007, 40(8).
- [3] KARP S. Google Ad Words: A Brief History Of Online Advertising Innovation[J]., 2008, 2.
- [4] POST T W. Google ends its dispute with yahoo.[Z]. 2004.
- [5] ANAGNOSTOPOULOS A, BRODER A Z, GABRILOVICH E, et al. Just-in-time contextual advertising[C]//Proceedings of the sixteenth ACM conference on Conference on information and knowledge management. 2007: 331-340.
- [6] BRODER A, FONTOURA M, JOSIFOVSKI V, et al. A semantic approach to contextual advertising[C]//Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. 2007: 559-566.
- [7] YUAN S, WANG J, ZHAO X. Real-time bidding for online advertising: measurement and analysis[C]//Proceedings of the Seventh International Workshop on Data Mining for Online Advertising. 2013: 3.
- [8] YUAN S, ABIDIN A Z, SLOAN M, et al. Internet advertising: An interplay among advertisers, online publishers, ad exchanges and web users[J]. ArXiv preprint arXiv:1206.1754, 2012.
- [9] ZHANG W, YUAN S, WANG J. Optimal real-time bidding for display advertising[C]// Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014: 1077-1086.
- [10] Fiksu Mobile DSP[Z]. <https://fiksu.com/>. Accessed: 2019-02.
- [11] SHEN J, ORTEN B, GEYIK S C, et al. From 0.5 million to 2.5 million: Efficiently scaling up real-time bidding[C]//Data Mining (ICDM), 2015 IEEE International Conference on. 2015: 973-978.

- [12] The Shanghai Stock Exchange Data[Z]. <http://www.sse.com.cn/market/stockdata/overview/day/>. Accessed: 2019-02.
- [13] WANG J, ZHANG W, YUAN S, et al. Display advertising with real-time bidding (RTB) and behavioural targeting[J]. *Foundations and Trends® in Information Retrieval*, 2017, 11(4-5): 297-435.
- [14] ZHANG W, YUAN S, WANG J, et al. Real-time bidding benchmarking with ipinyou dataset[J]. *ArXiv preprint arXiv:1407.7073*, 2014.
- [15] QU Y, CAI H, REN K, et al. Product-based neural networks for user response prediction[C]//*Data Mining (ICDM), 2016 IEEE 16th International Conference on*. 2016: 1149-1154.
- [16] REN K, ZHANG W, RONG Y, et al. User response learning for directly optimizing campaign performance in display advertising[C]//*Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. 2016: 679-688.
- [17] HE X, PAN J, JIN O, et al. Practical lessons from predicting clicks on ads at facebook[C]//*Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*. 2014: 1-9.
- [18] LEE K C, ORTEN B, DASDAN A, et al. Estimating conversion rate in display advertising from past performance data[C]//*Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2012: 768-776.
- [19] RENDLE S. Factorization machines[C]//*Data Mining (ICDM), 2010 IEEE 10th International Conference on*. 2010: 995-1000.
- [20] CAI H, REN K, ZHANG W, et al. Real-time bidding by reinforcement learning in display advertising[C]//*Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. 2017: 661-670.
- [21] WANG X, YU L, REN K, et al. Dynamic attention deep model for article recommendation by learning human editors' demonstration[C]//*Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2017: 2051-2059.
- [22] YUAN S, WANG J. Sequential selection of correlated ads by POMDPs[C]//*Proceedings of the 21st ACM international conference on Information and knowledge management*. 2012: 515-524.

- [23] REN K, ZHANG W, CHANG K, et al. Bidding Machine: Learning to Bid for Directly Optimizing Profits in Display Advertising[J]. IEEE Transactions on Knowledge and Data Engineering, 2018, 30(4): 645-659.
- [24] CUI Y, ZHANG R, LI W, et al. Bid landscape forecasting in online ad exchange marketplace[C]//Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. 2011: 265-273.
- [25] WU W C H, YE H M Y, CHEN M S. Predicting winning price in real time bidding with censored data[C]//Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2015: 1305-1314.
- [26] AMIN K, KEARNS M, KEY P, et al. Budget optimization for sponsored search: Censored learning in MDPs[J]. ArXiv preprint arXiv:1210.4847, 2012.
- [27] LIN C C, CHUANG K T, WU W C H, et al. Combining powers of two predictors in optimizing real-time bidding strategy under constrained budget[C]//Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. 2016: 2143-2148.
- [28] WU W, YE H M Y, CHEN M S. Deep censored learning of the winning price in the real time bidding[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018: 2526-2535.
- [29] WANG Y, REN K, ZHANG W, et al. Functional bid landscape forecasting for display advertising[C]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. 2016: 115-131.
- [30] ZHU W Y, SHIH W Y, LEE Y H, et al. A gamma-based regression for winning price estimation in real-time bidding advertising[C]//Big Data (Big Data), 2017 IEEE International Conference on. 2017: 1610-1619.
- [31] ZHANG W, ZHOU T, WANG J, et al. Bid-aware gradient descent for unbiased learning with censored data in display advertising[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016: 665-674.
- [32] KAPLAN E L, MEIER P. Nonparametric estimation from incomplete observations[J]. Journal of the American statistical association, 1958, 53(282): 457-481.
- [33] COX D R. Regression models and life-tables[J]. Journal of the Royal Statistical Society: Series B (Methodological), 1972, 34(2): 187-202.

- [34] ZHANG H H, LU W. Adaptive Lasso for Cox's proportional hazards model[J]. *Biometrika*, 2007, 94(3): 691-703.
- [35] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//*Advances in neural information processing systems*. 2012: 1097-1105.
- [36] GRAVES A, MOHAMED A R, HINTON G. Speech recognition with deep recurrent neural networks[C]//*Acoustics, speech and signal processing (icassp), 2013 IEEE international conference on*. 2013: 6645-6649.
- [37] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural computation*, 1997, 9(8): 1735-1780.
- [38] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[J]. *International Conference on Learning Representations*, 2015.
- [39] KATZMAN J L, SHAHAM U, CLONINGER A, et al. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network[J]. *BMC medical research methodology*, 2018, 18(1): 24.
- [40] ZHANG S, WU Y, CHE T, et al. Architectural complexity measures of recurrent neural networks[C]//*Advances in Neural Information Processing Systems*. 2016: 1822-1830.
- [41] WU S, REN W, YU C, et al. Personal recommendation using deep recurrent neural networks in NetEase[C]//*Data Engineering (ICDE), 2016 IEEE 32nd International Conference on*. 2016: 1218-1229.
- [42] ZHANG Y, DAI H, XU C, et al. Sequential Click Prediction for Sponsored Search with Recurrent Neural Networks.[C]//*AAAI:vol. 14*. 2014: 1369-1375.
- [43] ZHOU G, MOU N, FAN Y, et al. Deep Interest Evolution Network for Click-Through Rate Prediction[J]. *AAAI*, 2019.
- [44] WU Y, SCHUSTER M, CHEN Z, et al. Google's neural machine translation system: Bridging the gap between human and machine translation[J]. *ArXiv preprint arXiv:1609.08144*, 2016.
- [45] LIAO H, PENG L, LIU Z, et al. IPinYou global rtb bidding algorithm competition dataset[C]//*Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*. 2014: 1-6.

- [46] MASON S J, GRAHAM N E. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation[J]. Quarterly Journal of the Royal Meteorological Society, 2002, 128(584): 2145-2166.
- [47] BHATTACHARYA B, HABTZGHI D. Median of the p value under the alternative hypothesis[J]. The American Statistician, 2002, 56(3): 202-206.
- [48] LI Y, WANG J, YE J, et al. A multi-task learning formulation for survival analysis[C]// Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016: 1715-1724.
- [49] LEE C, ZAME W R, YOON J, et al. Deephit: A deep learning approach to survival analysis with competing risks[C]//. 2018.
- [50] RICHARDSON M, DOMINOWSKA E, RAGNO R. Predicting clicks: estimating the click-through rate for new ads[C]//Proceedings of the 16th international conference on World Wide Web. 2007: 521-530.
- [51] MCMAHAN H B, HOLT G, SCULLEY D, et al. Ad click prediction: a view from the trenches[C]//Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. 2013: 1222-1230.
- [52] OENTARYO R J, LIM E P, LOW J W, et al. Predicting response in mobile advertising with hierarchical importance-aware factorization machine[C]//Proceedings of the 7th ACM international conference on Web search and data mining. 2014: 123-132.
- [53] BERGER J O. Statistical decision theory and Bayesian analysis[M]. Springer Science & Business Media, 2013.
- [54] KRISHNA V. Auction theory[M]. Academic press, 2009.
- [55] WANG J, YUAN S. Real-time bidding: A new frontier of computational advertising research[C]//Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. 2015: 415-416.
- [56] MENON A K, CHITRAPURA K P, GARG S, et al. Response prediction using collaborative filtering with hierarchies and side-information[C]//Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. 2011: 141-149.

- [57] AGARWAL D, AGRAWAL R, KHANNA R, et al. Estimating rates of rare events with multiple hierarchies through scalable log-linear models[C]//Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. 2010: 213-222.
- [58] TA A P. Factorization machines with follow-the-regularized-leader for CTR prediction in display advertising[C]//Big Data (Big Data), 2015 IEEE International Conference on. 2015: 2889-2891.
- [59] LIU Q, YU F, WU S, et al. A convolutional click prediction model[C]//Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. 2015: 1743-1746.
- [60] CHAPELLE O. Offline evaluation of response prediction in online advertising auctions[C]//Proceedings of the 24th International Conference on World Wide Web. 2015: 919-922.
- [61] VASILE F, LEFORTIER D, CHAPELLE O. Cost-sensitive learning for utility optimization in online advertising auctions[C]//Proceedings of the ADKDD'17. 2017: 8.
- [62] EDELMAN B, OSTROVSKY M, SCHWARZ M. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords[J]. American economic review, 2007, 97(1): 242-259.
- [63] PERLICH C, DALESSANDRO B, HOOK R, et al. Bid optimizing and inventory scoring in targeted online advertising[C]//Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. 2012: 804-812.
- [64] XU J, SHAO X, MA J, et al. Lift-Based Bidding in Ad Selection.[C]//AAAI. 2016: 651-657.
- [65] ZHANG W, WANG J. Statistical arbitrage mining for display advertising[C]//Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2015: 1465-1474.
- [66] FRIEDMAN J, HASTIE T, TIBSHIRANI R, et al. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors)[J]. The annals of statistics, 2000, 28(2): 337-407.

- [67] DUPRET G E, PIWOWARSKI B. A user browsing model to predict search engine click data from past observations.[C]//Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval. 2008: 331-338.
- [68] JIANG M, CUI P, WANG F, et al. Fema: flexible evolutionary multi-faceted analysis for dynamic behavioral pattern discovery[C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014: 1186-1195.
- [69] KOREN Y. Factorization meets the neighborhood: a multifaceted collaborative filtering model[C]//Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. 2008: 426-434.
- [70] KOREN Y. Collaborative filtering with temporal dynamics[C]//Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. 2009: 447-456.
- [71] ZHOU G, ZHU X, SONG C, et al. Deep interest network for click-through rate prediction[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018: 1059-1068.
- [72] HIDASI B, KARATZOGLOU A. Recurrent neural networks with top-k gains for session-based recommendations[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management. 2018: 843-852.
- [73] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. Computer, 2009(8): 30-37.
- [74] ZHANG S, TAY Y, YAO L, et al. Next Item Recommendation with Self-Attention[J]. ArXiv preprint arXiv:1808.06414, 2018.
- [75] YING H, ZHUANG F, ZHANG F, et al. Sequential recommender system based on hierarchical attention networks[C]//The 27th International Joint Conference on Artificial Intelligence. 2018.
- [76] HIDASI B, KARATZOGLOU A, BALTRUNAS L, et al. Session-based recommendations with recurrent neural networks[J]. ArXiv preprint arXiv:1511.06939, 2015.
- [77] ZHENG L, NOROOZI V, YU P S. Joint deep modeling of users and items using reviews for recommendation[C]//Proceedings of the Tenth ACM International Conference on Web Search and Data Mining. 2017: 425-434.

- [78] REN K, FANG Y, ZHANG W, et al. Learning Multi-touch Conversion Attribution with Dual-attention Mechanisms for Online Advertising[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management. 2018: 1433-1442.
- [79] VILLATEL K, SMIRNOVA E, MARY J, et al. Recurrent Neural Networks for Long and Short-Term Sequential Recommendation[J]. ArXiv preprint arXiv:1807.09142, 2018.
- [80] HE R, MCAULEY J. Fusing similarity models with markov chains for sparse sequential recommendation[C]//2016 IEEE 16th International Conference on Data Mining (ICDM). 2016: 191-200.
- [81] HE R, FANG C, WANG Z, et al. Vista: A visually, socially, and temporally-aware model for artistic recommendation[C]//Proceedings of the 10th ACM Conference on Recommender Systems. 2016: 309-316.
- [82] WU C Y, AHMED A, BEUTEL A, et al. Recurrent recommender networks[C]//Proceedings of the tenth ACM international conference on web search and data mining. 2017: 495-503.
- [83] JING H, SMOLA A J. Neural survival recommender[C]//Proceedings of the Tenth ACM International Conference on Web Search and Data Mining. 2017: 515-524.
- [84] LIU Q, WU S, WANG D, et al. Context-aware sequential recommendation[C]//2016 IEEE 16th International Conference on Data Mining (ICDM). 2016: 1053-1058.
- [85] BEUTEL A, COVINGTON P, JAIN S, et al. Latent cross: Making use of context in recurrent recommender systems[C]//Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. 2018: 46-54.
- [86] TANG J, WANG K. Personalized top-n sequential recommendation via convolutional sequence embedding[C]//Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. 2018: 565-573.
- [87] KANG W C, MCAULEY J. Self-Attentive Sequential Recommendation[C]//2018 IEEE International Conference on Data Mining (ICDM). 2018: 197-206.
- [88] WESTON J, CHOPRA S, BORDES A. Memory networks[J]. ICLR, 2015.
- [89] SUKHBAATAR S, WESTON J, FERGUS R, et al. End-to-end memory networks[C]//Advances in neural information processing systems. 2015: 2440-2448.

- [90] KUMAR A, IRSOY O, ONDRUSKA P, et al. Ask me anything: Dynamic memory networks for natural language processing[C]//International Conference on Machine Learning. 2016: 1378-1387.
- [91] GRAVES A, WAYNE G, DANIHELKA I. Neural turing machines[J]. ArXiv preprint arXiv:1410.5401, 2014.
- [92] EBESU T, SHEN B, FANG Y. Collaborative memory network for recommendation systems[C]//The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. 2018: 515-524.
- [93] CHEN X, XU H, ZHANG Y, et al. Sequential recommendation with user memory networks[C]//Proceedings of the eleventh ACM international conference on web search and data mining. 2018: 108-116.
- [94] HUANG J, ZHAO W X, DOU H, et al. Improving sequential recommendation with knowledge-enhanced memory networks[C]//The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. 2018: 505-514.
- [95] WANG Q, YIN H, HU Z, et al. Neural memory streaming recommender networks with adversarial training[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018: 2467-2475.
- [96] SODHANI S, CHANDAR S, BENGIO Y. On Training Recurrent Neural Networks for Lifelong Learning[J]. ArXiv preprint arXiv:1811.07017, 2018.
- [97] MILLER A, FISCH A, DODGE J, et al. Key-value memory networks for directly reading documents[J]. ArXiv preprint arXiv:1606.03126, 2016.
- [98] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep learning[M]. MIT press, 2016.
- [99] KOUTNIK J, GREFF K, GOMEZ F, et al. A clockwork rnn[J]. ICML, 2014.
- [100] COGSWELL M, AHMED F, GIRSHICK R, et al. Reducing overfitting in deep networks by decorrelating representations[J]. ArXiv preprint arXiv:1511.06068, 2015.
- [101] CHEN Z, LIU B. Lifelong machine learning[J]. Synthesis Lectures on Artificial Intelligence and Machine Learning, 2016.
- [102] CHEN Z, HRUSCHKA JR E R, LIU B. Lifelong machine learning and computer reading the web[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016: 2117-2118.

- [103] MCAULEY J, TARGETT C, SHI Q, et al. Image-based recommendations on styles and substitutes[C]//Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2015: 43-52.
- [104] ZHU H, LI X, ZHANG P, et al. Learning Tree-based Deep Model for Recommender Systems[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018: 1079-1088.
- [105] ZHANG W, DU T, WANG J. Deep learning over multi-field categorical data[C]//European conference on information retrieval. 2016: 45-57.
- [106] AGARWAL D, GHOSH S, WEI K, et al. Budget pacing for targeted online advertisements at linkedin[C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014: 1613-1619.
- [107] LEE K C, JALALI A, DASDAN A. Real time bid optimization with smooth budget delivery in online advertising[C]//Proceedings of the Seventh International Workshop on Data Mining for Online Advertising. 2013: 1.
- [108] DIEMERT E, MEYNET J, GALLAND P, et al. Attribution Modeling Increases Efficiency of Bidding in Display Advertising[C]//Proceedings of the ADKDD'17. 2017: 2.
- [109] GEYIK S C, SAXENA A, DASDAN A. Multi-touch attribution based budget allocation in online advertising[C]//Proceedings of the Eighth International Workshop on Data Mining for Online Advertising. 2014: 1-9.
- [110] CHANDLER-PEPELNJAK J. Measuring roi beyond the last ad[J]. Atlas Institute Digital Marketing Insight, 2009: 1-6.
- [111] SHAO X, LI L. Data-driven multi-touch attribution models[C]//Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. 2011: 258-264.
- [112] DALESSANDRO B, PERLICH C, STITELMAN O, et al. Causally motivated attribution for online advertising[C]//Proceedings of the Sixth International Workshop on Data Mining for Online Advertising and Internet Economy. 2012: 7.
- [113] ZHANG Y, WEI Y, REN J. Multi-touch attribution in online advertising with survival theory[C]//2014 IEEE International Conference on Data Mining. 2014: 687-696.

- [114] JI W, WANG X. Additional multi-touch attribution for online advertising[C]//Thirty-First AAAI Conference on Artificial Intelligence. 2017.
- [115] XU L, DUAN J A, WHINSTON A. Path to purchase: A mutually exciting point process model for online advertising and conversion[J]. *Management Science*, 2014, 60(6): 1392-1412.
- [116] SONG J, XIAO J, WU F, et al. Hierarchical contextual attention recurrent neural network for map query suggestion[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2017, 29(9): 1888-1901.
- [117] JI W, WANG X, ZHANG D. A probabilistic multi-touch attribution model for online advertising[C]//Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. 2016: 1373-1382.
- [118] BERMAN R. Beyond the last touch: Attribution in online advertising[J]. *Marketing Science*, 2018, 37(5): 771-792.
- [119] SINHA R, SAINI S, ANADHAVELU N. Estimating the incremental effects of interactions for marketing attribution[C]//2014 International Conference on Behavioral, Economic, and Socio-Cultural Computing (BESC2014). 2014: 1-6.
- [120] YAN J, ZHANG C, ZHA H, et al. On machine learning towards predictive sales pipeline analytics[C]//Twenty-ninth AAAI conference on artificial intelligence. 2015.
- [121] YUAN S. Supply side optimisation in online display advertising[D]. UCL (University College London), 2015.
- [122] QIN Y, SONG D, CHENG H, et al. A dual-stage attention-based recurrent neural network for time series prediction[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. 2017: 2627-2633.
- [123] ZHAI S, CHANG K H, ZHANG R, et al. Deepintent: Learning attentions for online advertising with recurrent neural networks[C]//Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 2016: 1295-1304.
- [124] ROSENBLATT F. The perceptron: a probabilistic model for information storage and organization in the brain.[J]. *Psychological review*, 1958, 65(6): 386.

- [125] MA X, ZHAO L, HUANG G, et al. Entire Space Multi-Task Model: An Effective Approach for Estimating Post-Click Conversion Rate[C]//The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. 2018: 1137-1140.
- [126] GEHRING J, AULI M, GRANGIER D, et al. Convolutional sequence to sequence learning[C]//Proceedings of the 34th International Conference on Machine Learning-Volume 70. 2017: 1243-1252.
- [127] CAO X, CHEN H, WANG X, et al. Neural link prediction over aligned networks[C]//Thirty-Second AAAI Conference on Artificial Intelligence. 2018.
- [128] DUCHI J, SINGER Y. Efficient online and batch learning using forward backward splitting[J]. Journal of Machine Learning Research, 2009, 10(Dec): 2899-2934.
- [129] XIAO L. Dual averaging methods for regularized stochastic learning and online optimization[J]. Journal of Machine Learning Research, 2010, 11(Oct): 2543-2596.
- [130] ZHANG W. Optimal real-time bidding for display advertising[D]. UCL (University College London), 2016.
- [131] MENEZES F M, MONTEIRO P K. An introduction to auction theory[M]. Oxford University Press, USA, 2005.
- [132] KUHN S. Prisoner's Dilemma[J]. Stanford Encyclopedia of Philosophy, 2008.
- [133] RAPOPORT A, CHAMMAH A M. Prisoner's dilemma: A study in conflict and cooperation[M]. University of Michigan press, 1965.

## 致 谢

转眼间博士五年已经接近尾声，我也从人工智能的小白成长为略有贡献的入门者。这篇博士论文的内容涵盖了至少五篇国际会议长论文的内容，涉及的内容也包括机器学习、博弈论、优化等方面的知识。在这五年中，我从一位惊叹顶会论文模型之新奇、推导之严密的学术新人，成长为一名能有自己的学术见地的思考者。

在这段弥足珍贵的成长经历中，我首先要感谢我的授业恩师俞勇教授。在我刚拿到推免名额时有过一段时间的犹豫，担心自身能力是否能够满足博士研究的要求，也对未来的不确定性充满了疑虑。是他告诉我“人生道路上要学会坚持自己最初的梦想，不要因为外因而改变自己，因为在路上有太多的诱惑，如果什么都不能坚持的话，最终会完全背离自己的初衷，可能后悔的还是自己”。在俞老师的教诲下，我逐渐找到了做科研的兴趣与研究的方法，也走上了适合自己的学术道路。

我还要感谢在我读博期间带我走进计算广告研究领域的张伟楠老师。他敏锐的科研嗅觉、踏实的研究态度与全面的研究思路不断影响着我。我的博士论文中的多个研究课题与涉及的多方面研究方向，均经过我和张老师的不断讨论、打磨。他是我科研道路上非常敬佩的一位学者，希望之后还能继续与张老师在人工智能研究中一起探索与前进。

科研与生活是一位科研人员人生道路的左右两边，我的爱人冯逸在生活中给予了我许多照顾，同时也一直在支持我的研究工作。在我读博期间，我们步入了婚姻殿堂，之后还有了一位可爱的小宝宝。我非常感谢她，希望在之后的工作与生活中一同奋斗，一同享受精彩。

我还要感谢所有与我有过合作的老师、同学，包括 UCL 教授汪军老师、美团广告团队算法工程师容一飞，以及朱晨浩、黄伟岳、周志明、邱霖、曹雪智、蔡涵、房宇辰、秦佳锐、郑雷等同学。他们对我的研究论文的修改意见与建议弥足珍贵，同时又给予了我研究思路上的众多启迪。最后我要感谢我的父母，他们在我攻读博士期间无私地为我提供经济上与精神上的支持，没有他们的陪伴与鼓励，我也无法获得自身的成长与突破。

## 攻读学位期间发表的学术论文

- [1] REN K, QIN J R, ZHENG LEI, et al. Deep Landscape Forecasting for Real-time Bidding Advertising[C]. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2019.
- [2] REN K, QIN J R, FANG Y C, et al. Lifelong Sequential Modeling with Personalized Memorization for User Response Prediction[C]. Proceedings of 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2019).
- [3] REN K, QIN J R, et al. Deep Recurrent Survival Analysis[C]. Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19).
- [4] REN K, FANG Y C, ZHANG W N, et al. Learning Conversion Attribution with Dual-attention Mechanism for Online Advertising[C]. Proceedings of the 2018 ACM on Conference on Information and Knowledge Management, ACM, 2018: 1433-1442.
- [5] REN K, ZHANG W N, et al. Bidding Machine: Learning to Bid for Directly Optimizing Profits in Display Advertising[J]. IEEE Transactions on Knowledge and Data Engineering, vol. 30, no. 4, pp. 645-659, April 1 2018.
- [6] REN K, ZHANG W N, et al. User Response Learning for Directly Optimizing Campaign Performance Display Advertising[C]. Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. ACM, 2016: 679-688.
- [7] SONG Y X, REN K, CAI H, ZHANG W N. Volume Ranking and Sequential Selection Programmatic Display Advertising[C]. Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, ACM, 2017, pp. 1099-1107.
- [8] WANG X J, YU L T, REN K, ZHANG W N, et al. Dynamic Attention Deep Model for Article Recommendation by Learning Human Editors' Demonstration[C]. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2017, pp. 2051-2059.

- [9] CAI H, REN K, ZHANG W N, et al. Real-Time Bidding by Reinforcement Learning in Display Advertising[C]. Proceedings of the Tenth ACM International Conference on Web Search and Data Mining Pages 661-670.
- [10] QU Y R, CAI H, REN K, ZHANG W N, et al. Product-Based Neural Networks for User Response Prediction[C]. 2016 IEEE 16th International Conference on Data Mining (ICDM), Barcelona, 2016, pp. 1149-1154.
- [11] WANG Y C, REN K, ZHANG W N, et al. Functional Bid Landscape Forecasting for Display Advertising[C]. Joint European Conference on Machine Learning and Knowledge Discovery Databases, Springer, Cham, 2016, pp. 115-131.
- [12] ZHU C H, REN K, WANG H F, et al. A Graph Traversal Based Approach to Answer Non-Aggregation Questions over DBpedia[C]. Joint International Semantic Technology Conference, Springer, Cham, 2015, pp. 219-234.
- [13] LU G S, ZHOU Z M, SONG Y X, REN K, YU Y. Guiding the One-to-one Mapping in CycleGAN via Optimal Transport[C]. Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19).
- [14] ZHOU Z M, CAI H, SONG Y X, REN K, et al. Activation Maximization Generative Adversarial Nets[C]. International Conference on Learning Representations 2018.
- [15] ZHANG H F, ZHANG W N, RONG Y F, REN K, et al. Managing Risk of Bidding Display Advertising[C]. Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, ACM, 2017, pp. 581-590.

## 攻读学位期间参与的项目

- [1] 面上基金“基于深度神经网络的个性化推荐系统” (61772333)
- [2] 重点项目“网络广告拍卖的博弈理论和计算方法研究” (61632017)
- [3] 青年基金“基于多域离散数据的深度学习” (61702327)
- [4] 上海扬帆计划 (17YF1428200)