分类号 学号M201778277

学校代码 10487 密级



**硕士学位论文**

**点击率预估在推荐系统中的优化方法研究**

|  |  |
| --- | --- |
| 学位申请人： | 闫 昊 |
| 学科专业： | 软件工程 |
| 指导教师： | 邱德红 教授 |
| 答辩日期： | 2019.11.28 |

**A Dissertation Submitted in Partial Fulfillment of the Requirements**

**for the Degree of Master of Engineering**

**Research on optimization method of click rate estimation in recommendation system**

|  |  |
| --- | --- |
| **Candidate :** | **Yan Hao** |
| **Major :** | **Software Engeering** |
| **Supervisor :** | **Prof. Qiu Dehong** |
|  |  |

**Huazhong University of Science & Technology**

**Wuhan 430074, P. R. China**

**May, 2019**

独创性声明

本人声明所呈交的学位论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除文中已经标明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对本文的研究做出贡献的个人和集体，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：闫昊

日期： 2019年 5月 26日

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，即：学校有权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。本人授权华中科技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。

保密□， 在 年解密后适用本授权书。

本论文属于

不保密□√。

（请在以上方框内打“√”）

学位论文作者签名： 闫昊 指导教师签名：

日期： 2019年 5月 26日 日期： 年 月 日

# 摘 要

点击率预估在目前的计算广告和推荐系统中都得到了广泛的应用，并且广告业务也是大部分公司的核心变现业务，因此CTR预估得到了越来越多企业的重视和应用研究。

近些年来，深度学习发展迅速，业界对其的应用也日渐成熟，深度学习一个极大的优点就是能够很好的处理非线性相关问题，能够挖掘出普通机器学习算法无法挖掘出的特征组合序列，并且在真实数据测试的结果上表现优异。

业界目前在将深度学习方法应用于CTR预估也有不少成果，如Wide&Deep,DNN,FM但局限于目前深度学习框架的多样性，没有一个统一的理论，但目的都是致力于使用新的框架，通过实验的方法来检测是否能够提高CTR预估准确率。

因此，结合我在公司实习这段时期内接触到公司内部的CTR预估模型，我希望能够实现CTR预估在推荐系统中的优化方法研究，并致力于研究出能够解决目前在工作中遇到的问题的新方法，来减少公司在该领域的投入人力并提高CTR预估准确率,并跟上深度学习框架快速的发展，提高部门内部的推荐系统中CTR排序的准确性。

**关键词：** 点击率预估 深度学习 优化方法研究 推荐系统

# ABSTRACT

The click rate prediction has been widely used in the current computing advertising and recommendation systems, and the advertising business is also the core realizing business of most companies. Therefore, the CTR estimation has received more and more enterprise attention and application research.

In recent years, deep learning has developed rapidly, and the application of the industry has become more and more mature. A great advantage of deep learning is that it can deal with nonlinear related problems well, and can mine the feature combination sequence that ordinary machine learning algorithms cannot mine. And excellent in the results of real data testing.

The industry is currently applying a deep learning method to CTR estimates. There are also many achievements, such as Wide&Deep, DNN, and FM, but limited to the diversity of current deep learning frameworks. There is no unified theory, but the goal is to use new frameworks. , through experimental methods to detect whether it can improve the accuracy of CTR estimation.

Therefore, in combination with my internal CTR estimation model during the internship period of the company, I hope to be able to achieve CTR estimation in the optimization method of the recommendation system, and work hard to solve the problem that can be solved at present. A new approach to the problem, to reduce the company's investment in the field and improve the accuracy of CTR estimates, and to keep up with the rapid development of the deep learning framework, improve the accuracy of CTR sequencing in the department's internal recommendation system.

**Key words： Click rate prediction, deep learning, optimization method research, recommendation system**

目 录

[**摘 要 4**](#_Toc9870023)

[**ABSTRACT 5**](#_Toc9870024)

[**一 点击率预估理论概论 7**](#_Toc9870025)

[**1.1 点击率预估的概念及商业应用 7**](#_Toc9870026)

[**1.2 点击率预估当前的发展情况概述 7**](#_Toc9870027)

[**1.3 点击率预估当前的发展情况概述 9**](#_Toc9870028)

[**二 点击率预估在当前公司的应用 10**](#_Toc9870029)

[**2.1 利用LR进行稀疏场景下的CTR测试及其存在的问题 10**](#_Toc9870030)

[**2.2 利用GBDT+LR进行CTR测试及其存在的问题 10**](#_Toc9870031)

[**2.3 利用FM进行CTR测试及其存在的问题 10**](#_Toc9870032)

[**2.4 利用DeepFM进行CTR测试及其存在的问题 10**](#_Toc9870033)

[**三 点击率预估的优化方法研究 10**](#_Toc9870034)

[**3.1 深度学习在点击率预估中的应用分析 10**](#_Toc9870035)

[**3.2 点击率预估的优化方法研究过程 10**](#_Toc9870036)

[**3.3 点击率预估的优化方法研究结果测试与分析 10**](#_Toc9870037)

[**3.4 对比第二章中的实验结果并总结 10**](#_Toc9870038)

[**参考文献 11**](#_Toc9870039)

# 一 点击率预估理论概论

随着现代科学技术的不断进步，点击率预估在目前的计算广告和推荐系统中都得到了广泛的应用，并且广告业务也是大部分公司的核心变现业务，因此CTR预估得到了越来越多企业的重视和应用研究。

## 1.1 点击率预估的概念及商业应用

点击率预估，顾名思义，即对用户对某个事务进行点击的概率大小的预测，在现在的互联网科技公司里面有着非常重要的作用，我们知道大部分互联网公司的相当一收入都来自于广告业务，业界也经常流传着一些故事，某某科学家通过建立更好的点击率预测模型，为公司带来了上亿的增量收入，由此可见点击率预估作为公司的核心业务一点也不为过。

点击率预估要求公司技术人员根据公司产品的定位以及应用场景，给用户推送适当的广告，并使得用户点击广告的概率最大，这样对于公司来说可以提高广告业务的收入，对于广告主来说，自己投放的广告也尽可能大概率的的得到了用户的点击，那么点击过广告的部分用户就可能会购买该公司的产品，由此看来，提高点击率可以实现双方的共赢，因此，这也是点击率预估近些年来发展迅速的一大原因之一。

## 1.2 点击率预估当前的发展情况概述

CTR即点击率预估，在目前的计算广告[[1]](#footnote-1)[1]和推荐系统中都得到了广泛的应用，并且广告业务也是大部分公司的核心变现业务，因此CTR预估得到了越来越多企业的重视和应用研究。

CTR预估中用的最多的模型是LR[[2]](#footnote-2)[4]（Logistic Regression），LR是广义线性模型，与传统线性模型相比，LR使用了Logit变换将函数值映射到0~1区间，映射后的函数值就是CTR的预估值。LR这种线性模型很容易并行化，处理上亿条训练样本不是问题，但线性模型学习能力有限，需要大量特征工程预先分析出有效的特征、特征组合，从而去间接增强LR的非线性学习能力。

LR模型中的特征组合很关键， 但又无法直接通过特征笛卡尔积解决，只能依靠人工经验，耗时耗力同时并不一定会带来效果提升。如何自动发现有效的特征、特征组合，弥补人工经验不足，缩短LR特征实验周期，是亟需解决的问题。Facebook 2014年的文章介绍了通过GBDT[[3]](#footnote-3)[24]（Gradient Boost Decision Tree）解决LR的特征组合问题，随后Kaggle竞赛也有实践此思路，GBDT与LR融合开始引起了业界关注。

随着CTR预估技术的发展，Facebook在2014年发表了“Practical Lessons from Predicting Clicks on Ads at Facebook”[[4]](#footnote-4)[41], 在这篇文章中，Facebook 提出了经典的 GBDT （Gradient Boosting Decision Trees）+LR(Logistics Regression) 的 CTR 模型结构，可以说开启了特征工程模型化、自动化的新阶段。

简而言之，文章提出了一种利用 GBDT 自动进行特征筛选和组合，进而生成新的 feature vector，再把该 feature vector 当作 logisticre gression 的模型输入，预测 CTR 的模型结构。

GBDT算法的特点正好可以用来发掘有区分度的特征、特征组合，减少特征工程中人力成本，且计算广告[[5]](#footnote-5)[1]业界现在已有实践。

GBDT（Gradient Boost Decision Tree）是一种常用的非线性模型，它基于集成学习中的boosting思想，每次迭代都在减少残差的梯度方向新建立一颗决策树，迭代多少次就会生成多少颗决策树。GBDT的思想使其具有天然优势可以发现多种有区分性的特征以及特征组合，决策树的路径可以直接作为LR输入特征使用，省去了人工寻找特征、特征组合的步骤。使用GBDT+LR前后的特征，融合前人工寻找有区分性特征（raw feature）、特征组合（cross feature），融合后直接通过黑盒子（Tree模型GBDT）进行特征、特种组合的自动发现。

GBDT与LR的融合方式，在上述的Facebook paper中有个例子Tree1、Tree2为通过GBDT模型学出来的两颗树，x为一条输入样本，遍历两棵树后，x样本分别落到两颗树的叶子节点上，每个叶子节点对应LR一维特征，那么通过遍历树，就得到了该样本对应的所有LR特征。由于树的每条路径，是通过最小化均方差等方法最终分割出来的有区分性路径，根据该路径得到的特征、特征组合都相对有区分性，效果理论上不会亚于人工经验的处理方式。

GBDT模型的特点，非常适合用来挖掘有效的特征、特征组合。业界不仅GBDT+LR融合有实践，GBDT+FM也有实践，2014 Kaggle CTR竞赛冠军就是使用GBDT+FM，2016年Youtube提出的DNN用于推荐系统论文中提出的深度学习模型[[6]](#footnote-6)[37]，可见，使用GBDT融合其它模型是非常有效果的尝试。

通过GBDT生成的特征，可直接作为LR的特征使用，省去人工处理分析特征的环节，LR的输入特征完全依赖于通过GBDT得到的特征。业界已经通过实验发现GBDT+LR在曝光充分的广告上确实有效果，但整体效果需要权衡优化各类树的使用。同时，也可考虑将GBDT生成特征与LR原有特征结合起来使用。

## 1.3 点击率预估当前的发展情况概述

从上述的介绍中，可以看出传统机器学习方法如LR，GBDT等算法在CTR预估中能够起到重要的作用，但是不足之处也很明显。点击率预估模型涉及的训练样本一般是上亿级别，样本量大，模型常采用速度较快的LR。但LR是线性模型，学习能力有限，此时特征工程尤其重要。现有的特征工程实验，主要集中在寻找到有区分度的特征、特征组合，折腾一圈未必会带来效果提升。

为了解决这一问题，计算广告[[7]](#footnote-7)[13,14]业内通过使用GBDT来进行自动意义上的特征选择，GBDT 是由多棵回归树组成的树林，后一棵树利用前面树林的结果与真实结果的残差做为拟合目标。每棵树生成的过程是一棵标准的回归树生成过程，因此每个节点的分裂是一个自然的特征选择的过程，而多层节点的结构自然进行了有效的特征组合，也就非常高效的解决了过去非常棘手的特征选择和特征组合的问题。

目前业界也有一些效果不错的非线性模型不断被提出来，并被工程实践且取得不错效果，但这些模型都或多或少存在一些不足。比如Kernel方法，因为复杂度太高而不易实现；比如Tree based方法，这个是由Facebook团队在2014年首先提出，有效地解决了LR模型的特征组合问题，但缺点就是仍然是对历史行为的记忆，缺乏推广性；还有FM[[8]](#footnote-8)[32]（factorization machine）模型，能自动学习高阶属性的权值，不用通过人工的方式选取特征来做交叉，但FM模型只能拟合特定的非线性模式，如最常用的2阶FM只能拟合特征之间的线性关系和二次关系。深度神经网络非线性拟合能力足够强，但面对广告这样的大规模工业级稀疏数据，适合数据规律的、具备推广性的网络结构业界依然在探索中，尤其是要做到端到端规模化上线，这里面的技术挑战依然很大。

随着大数据技术的发展，伴随着数据量一天天爆炸性增长，数据的维度也在不断增加，这给有效特征的挑选带来了极大的挑战。因此上述GBDT+LR,GBDT+FM的组合在CTR预估的准确率上也面临着挑战。

# 二 点击率预估在当前公司的应用

## 2.1 利用逻辑回归进行稀疏场景下的CTR测试及其存在的问题

### 2.1.1 逻辑回归算法介绍

Logistic regression（逻辑回归）是一种非线性回归模型，特征数据可以是连续的，也可以是分类变量和哑变量，是当前业界比较常用的[机器学习](http://www.dataguru.cn/article-4063-1.html?union_site=innerlink" \t "_blank)方法，用于估计某种事物的可能性，主要的用途：

分类问题：如，反垃圾系统判别，通过计算被标注为垃圾邮件的概率和非垃圾邮件的概率判定；

排序问题：如，[推荐系统](http://www.dataguru.cn/article-4537-1.html?union_site=innerlink)中的排序，根据转换预估值进行排序；

预测问题：如，广告系统中CTR预估，根据CTR预估值预测广告收益；

这个世界是随机的，所以万物的发生都可以用可能性或者几率（Odds）来表达。“几率”指的是某事物发生的可能性与不发生的可能性的比值。

LR可以用来回归，也可以用来分类，主要是二分类。logistic回归模型在多分类问题上的推广是softmax regression。

### 2.1.2 逻辑回归用于点击率预估试验测试

**//此处添加真实数据集下的测试结果！**

### 2.1.3 逻辑回归用于点击率预估的优缺点分析

优点：

1. 是一个很好的baseline，效果不错。

2. 实现简单，有开源的工具可以直接用来训练，在线的代码也写起来也比较容易。

3. 易于并行化处理。

缺点：

1. 因为是线性模型，所以有选择交叉特征的工作，这部分工作消耗大量的精力，但往往没什么效果。一般都是wrapper方法选择，每轮可能都要进行小时级的运算，理论上要进行2^n轮(n是特征数)，但因为离线分析的指标和线上效果不一定是强相关的，所以分析起来存在困难。再者点击率预估的数据变化是比较大的，离线选出来的特征，参数都不一定适用于未来，也就是说LR模型的泛化能力并不强。

2. 因为广告位对点击率有着决定性的影响，所以几乎所有的特征都会交叉广告位，这样广告位间的信息是无法共享的，比如广告A在广告位x上点击率很高，但如果广告A在广告位y上没有曝光过，那对广告位y来讲，广告A是一个全新的广告。

3. 实践中交叉特征还会导致模型非常大，如果模型非常大，会导致同步模型变慢，一样会严重影响效果。

4. 在线预测时，因为大量的特征都要与广告ID交叉，所以在线拼装特征的成本很高，性能可能也会成为问题。

## 2.2 利用GBDT+LR进行CTR测试及其存在的问题

### 2.2.1 GBDT+LR算法介绍

在CTR预估问题的发展初期，使用最多的方法就是逻辑回归(LR)，LR使用了Sigmoid变换将函数值映射到0~1区间，映射后的函数值就是CTR的预估值。  
LR属于线性模型，容易并行化，可以轻松处理上亿条数据。 但是学习能力十分有限，需要大量的特征工程来增加模型的学习能力。但大量的特征工程耗时耗力同时并不一定会带来效果提升。因此，如何自动发现有效的特征、特征组合，弥补人工经验不足，缩短LR特征实验周期，是亟需解决的问题。

Facebook 2014年的文章介绍了通过GBDT解决LR的特征组合问题，随后Kaggle竞赛也有实践此思路，GBDT与LR融合开始引起了业界关注。

在介绍这个模型之前，我们先来介绍两个问题：  
1）为什么要使用集成的决策树模型，而不是单棵的决策树模型：一棵树的表达能力很弱，不足以表达多个有区分性的特征组合，多棵树的表达能力更强一些。可以更好的发现有效的特征和特征组合。  
2）为什么建树采用GBDT而非RF（随机森林）：RF也是多棵树，但从效果上有实践证明不如GBDT。且GBDT前面的树，特征分裂主要体现对多数样本有区分度的特征；后面的树，主要体现的是经过前N颗树，残差仍然较大的少数样本。优先选用在整体上有区分度的特征，再选用针对少数样本有区分度的特征，思路更加合理，这应该也是用GBDT的原因。

### 2.2.2 GBDT+LR算法的融合方案

GBDT和LR的融合方案，FaceBook的paper中有个例子：

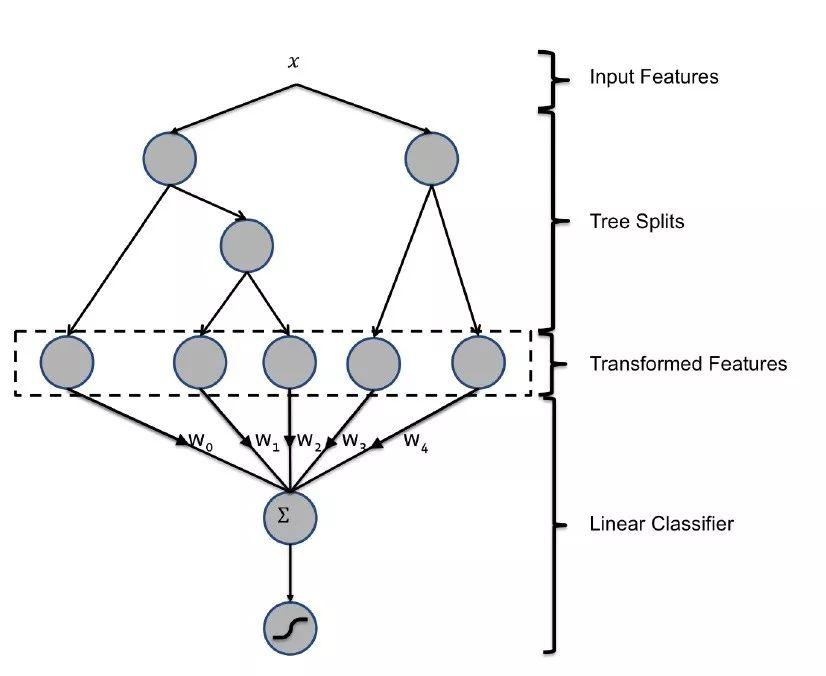


图1-GBDT+LR 模型结构

图中共有两棵树，x为一条输入样本，遍历两棵树后，x样本分别落到两颗树的叶子节点上，每个叶子节点对应LR一维特征，那么通过遍历树，就得到了该样本对应的所有LR特征。构造的新特征向量是取值0/1的。举例来说：上图有两棵树，左树有三个叶子节点，右树有两个叶子节点，最终的特征即为五维的向量。对于输入x，假设他落在左树第一个节点，编码[1,0,0]，落在右树第二个节点则编码[0,1]，所以整体的编码为[1,0,0,0,1]，这类编码作为特征，输入到LR中进行分类。

由于树的每条路径，是通过最小化均方差等方法最终分割出来的有区分性路径，根据该路径得到的特征、特征组合都相对有区分性，效果理论上不会亚于人工经验的处理方式。

## 2.3 利用FM进行CTR测试及其存在的问题

## 2.4 利用DeepFM进行CTR测试及其存在的问题

# 三 点击率预估的优化方法研究

## 3.1 深度学习在点击率预估中的应用分析

## 3.2 点击率预估的优化方法研究过程

## 3.3 点击率预估的优化方法研究结果测试与分析

## 3.4 对比第二章中的实验结果并总结

# 参考文献

[1] 计算广告[M]. 人民邮电出版社 , 刘鹏, 2015

[2] 统计学习方法[M]. 清华大学出版社 , 李航, 2012

[3] 广告点击率预估技术综述[J]. 浙江理工大学学报, 陈巧红,余仕敏,贾宇波. 2015(11).

[4] Logistic回归模型[M]. 高等教育出版社 , 王济川,郭志刚[著], 2001

[5] [基于大数据平台的精准广告系统研究与设计](https://kns.cnki.net/kcms/detail/detail.aspx?filename=DNXJ201504015&dbcode=CJFQ&dbname=CJFD2015&v=)[J]. 张建,孙铭,段娟.  电脑与信息技术. 2015(04)

[6] [计算广告:以数据为核心的Web综合应用](https://kns.cnki.net/kcms/detail/detail.aspx?filename=JSJX201110007&dbcode=CJFQ&dbname=CJFD2011&v=)[J]. 周傲英,周敏奇,宫学庆.  计算机学报. 2011(10)

[7] [电子商务推荐模型的用户特征提取方法及系统](https://kns.cnki.net/kcms/detail/detail.aspx?filename=CN104331817A&dbcode=SCPD&v=)[P]. 沈慧,范小朋,赵东辉,须成忠. 中国专利:CN104331817A,

[8] [大数据时代背景下的广告营销模式研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=1014297322.nh&dbcode=CMFD&dbname=CMFD2014&v=)[D]. 张之怡.苏州大学 2014

[9] [面向自然语言处理的深度学习研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=MOTO201610001&dbcode=CJFQ&dbname=CJFD2016&v=)[J]. 奚雪峰,周国栋.  自动化学报. 2016(10)

[10] [在线广告中高层特征表示及点击率预测方法研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=1014083837.nh&dbcode=CMFD&dbname=CMFD2015&v=)[D]. 邵迪.哈尔滨工业大学 2014

[11] [不同池化模型的卷积神经网络学习性能研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=ZGTB201609007&dbcode=CJFQ&dbname=CJFD2016&v=)[J]. 刘万军,梁雪剑,曲海成.  中国图象图形学报. 2016(09)

[12] [广告点击率预估的深层神经网络模型研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=1015583935.nh&dbcode=CMFD&dbname=CMFD2015&v=)[D]. 王孝舒.北京邮电大学 2015

[13] [基于特征学习的广告点击率预估技术研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=1017241252.nh&dbcode=CMFD&dbname=CMFD2018&v=)[D]. 周永.哈尔滨工程大学 2014

[14] [基于特征学习的广告点击率预估技术研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=JSJX201604010&dbcode=CJFQ&dbname=CJFD2016&v=)[J]. 张志强,周永,谢晓芹,潘海为.  计算机学报. 2016(04)

[15] [基于LDA的互联网广告点击率预测研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=JSYJ201604004&dbcode=CJFQ&dbname=CJFD2016&v=)[J]. 朱志北,李斌,刘学军,胡平.  计算机应用研究. 2016(04)

[16] 计算广告学将成为数字商业的奠基学科[J]. 李敏.  程序员. 2014 (05)

[17] [基于平衡采样的轻量级广告点击率预估方法](https://kns.cnki.net/kcms/detail/detail.aspx?filename=JSYJ201401007&dbcode=CJFQ&dbname=CJFD2014&v=)[J]. 施梦圜,顾津吉.  计算机应用研究. 2014(01)

[18] [广告点击率估算技术综述](https://kns.cnki.net/kcms/detail/detail.aspx?filename=HDSZ201303005&dbcode=CJFQ&dbname=CJFD2013&v=)[J]. 纪文迪,王晓玲,周傲英.  华东师范大学学报(自然科学版). 2013(03)

[19] [基于联合概率矩阵分解的上下文广告推荐算法](https://kns.cnki.net/kcms/detail/detail.aspx?filename=RJXB201303002&dbcode=CJFQ&dbname=CJFD2013&v=)[J]. 涂丹丹,舒承椿,余海燕.  软件学报. 2013(03)

[20] [基于贝叶斯网的广告点击率预测方法及实现](https://kns.cnki.net/kcms/detail/detail.aspx?filename=1013306802.nh&dbcode=CMFD&dbname=CMFD2014&v=)[D]. 王朝禄.云南大学 2013

[21] [网络广告投放算法的研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=1014183089.nh&dbcode=CMFD&dbname=CMFD2014&v=)[D]. 霍艳.东北大学 2013

[22] [大数据分析的无限深度神经网络方法](https://kns.cnki.net/kcms/detail/detail.aspx?filename=JFYZ201601007&dbcode=CJFQ&dbname=CJFD2016&v=)[J]. 张蕾,章毅.  计算机研究与发展. 2016(01)

[23] [面向大数据分析的在线学习算法综述](https://kns.cnki.net/kcms/detail/detail.aspx?filename=JFYZ201508002&dbcode=CJFQ&dbname=CJFD2015&v=)[J]. 李志杰,李元香,王峰,何国良,匡立.  计算机研究与发展. 2015(08)

[24] 数据挖掘中的新方法[M]. 科学出版社 , 邓乃扬,田英杰著, 2004

[25] [协同过滤推荐算法综述](https://kns.cnki.net/kcms/detail/detail.aspx?filename=XXWX200907006&dbcode=CJFQ&dbname=CJFD2009&v=)[J]. 马宏伟,张光卫,李鹏.  小型微型计算机系统. 2009(07)

[26] [一种基于逻辑回归模型的搜索广告点击率预估方法的研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=1013178340.nh&dbcode=CMFD&dbname=CMFD2014&v=)[D]. 王兵.浙江大学 2013

[27] [个性化广告点击率预测的研究和实现](https://kns.cnki.net/kcms/detail/detail.aspx?filename=1013241851.nh&dbcode=CMFD&dbname=CMFD2013&v=)[D]. 司向辉.北京邮电大学 2013

[28] Ying Shan, T Ryan Hoens, Jian Jiao, Haijing Wang, Dong Yu, and JC Mao. 2016. Deep Crossing: Web-Scale Modeling without Manually Cra ed Combinatorial Features. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 255–262.

[29] Wang, R., Fu, B., Fu, G., Wang, M.: Deep & cross network for ad click predictions. In: Proceedings of the ADKDD 17. pp. 12:1–12:7 (2017).

[30][Gai et al] Learning Piece-wise Linear Models from Large Scale Data for Ad Click Prediction

[31][Cheng et al., 2016] Wide & deep learning for recommender systems. CoRR.

[32][Huifeng Guo et al, 2017] DeepFM: A Factorization-Machine based Neural Network for CTR Prediction.

[33][Guorui Zhou et al, 2017] Deep Interest Network for Click-Through Rate Prediction.

[34][He X, Pan J, Jin O, et al, 2014] Practical lessons from predicting clicks on ads. at facebook. ACM SIGKDD.

[35][Rendle, 2010] Factorization machines. In ICDM.

[36] Practical Lessons from Predicting Clicks on Ads at Facebook. He,X.,et al. ADKDD’’14 . 2014

[37] Paul Covington, Jay Adams, and Emre Sargin. Deep neural networks for youtube recommendations. In Proceedings of the 10th ACM Conference on Recommender Systems, pages 191–198. ACM, 2016.

[38] Heng-Tze Cheng and Levent Koc. Wide & deep learning for recommender systems. In Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, pages 7–10. ACM, 2016.

[39] Scalable training of L1-regularized log-linear models. Andrew G,Gao J. Pro-ceedings of the24th International Conference on Machine Learning . 2007

[40] Dual Averaging Methods for Regularized Stochastic Learning and OnlineOptimization. Lin X. NIPS . 2009

[41] [[GBDT+LR] Practical Lessons from Predicting Clicks on Ads at Facebook (Facebook 2014)](https://github.com/wzhe06/Ad-papers/blob/master/Classic%20CTR%20Prediction/%5BGBDT%2BLR%5D%20Practical%20Lessons%20from%20Predicting%20Clicks%20on%20Ads%20at%20Facebook%20%28Facebook%202014%29.pdf)

[42] [XGBoost：A Scalable Tree Boosting System](https://arxiv.org/pdf/1603.02754.pdf)

1. [1] 计算广告[M]. 人民邮电出版社 , 刘鹏, 2015 [↑](#footnote-ref-1)
2. [4] Logistic回归模型[M]. 高等教育出版社 , 王济川,郭志刚[著], 2001 [↑](#footnote-ref-2)
3. [24] 统计学习方法[M]. 清华大学出版社 , 李航, 2012 [↑](#footnote-ref-3)
4. [41] [[GBDT+LR] Practical Lessons from Predicting Clicks on Ads at Facebook (Facebook 2014)](https://github.com/wzhe06/Ad-papers/blob/master/Classic%20CTR%20Prediction/%5BGBDT%2BLR%5D%20Practical%20Lessons%20from%20Predicting%20Clicks%20on%20Ads%20at%20Facebook%20%28Facebook%202014%29.pdf)  [↑](#footnote-ref-4)
5. [1] 计算广告[M]. 人民邮电出版社 , 刘鹏, 2015 [↑](#footnote-ref-5)
6. [37] Deep neural networks for youtube recommendations. [↑](#footnote-ref-6)
7. [13,14] [基于特征学习的广告点击率预估技术研究](https://kns.cnki.net/kcms/detail/detail.aspx?filename=1017241252.nh&dbcode=CMFD&dbname=CMFD2018&v=) [↑](#footnote-ref-7)
8. [32] [[FM] Fast Context-aware Recommendations with Factorization Machines (UKON 2011)](https://github.com/wzhe06/Ad-papers/blob/master/Classic%20CTR%20Prediction/%5BFM%5D%20Fast%20Context-aware%20Recommendations%20with%20Factorization%20Machines%20%28UKON%202011%29.pdf)  [↑](#footnote-ref-8)