

基于 LSTM 的商品推荐模型研究

张 洁

(东华大学 上海 201620)

摘 要 :与传统电商平台用户的购物特点有所不同,新零售平台的用户购买商品具有重复率高、购买周期短的特点,覆盖衣食住行各个方面,很多属于快销品。构建一个基于新零售电商平台的商品重构推荐系统,对用户的周期性购买行为进行分析预测,有利于提高用户购物体验,增加平台盈利。目前基于大数据的数据挖掘应用方面很多,近些年,深度学习在数据挖掘方面展露头角,深度学习具有多隐层的多层感知器,通过由低层到高层的属性特征分析来发现数据的分布式特征,这对商品推荐系统中分析用户行为具有传统分析方式所不具有的优势。

关键词 新零售;商品推荐;深度学习

中图分类号:TP391.3

文献标识码:A

文章编号:2096-4390(2021)11-0088-02

1 概述

电商行业的快速发展使得用户购买的商品由购买服装、日用品、电子产品等,转为购买肉类、蔬菜、水果、等生鲜类商品。不同生鲜类别的商品周期性规律不同。新零售平台需要对用户的周期性购买行为进行建模^[1]。但新零售平台用户量巨大,对所有用户的周期性购买行为进行建模具有较大难度。因此,商品重构推荐模型的构建特点是要着重挖掘商品购买记录中的时序信息。

2 相关概念

2.1 Keras

Keras 是由纯 python 编写的深度学习框架,是一个高层神经网络 API,支持快速实验,具有用户友好、模块化、易扩展的优点。Keras 没有单独的模型配置文件,模型由 python 代码表述,更紧凑和更易 debug。网络层、损失函数、优化器等都是独立的模块,可以使用它们构建自己的模型^[2]。根据需要想要添加新模块比较容易,只需要仿照现有的模块编写类和方法函数即可。模块创新的便利性使得 Keras 更加适合模型训练研究^[3]。

2.2 深度学习模型

2.2.1 RNN 模型

递归神经网络(也叫 RNN 算法)源自于 1982 年由 Sarath Sathasivam 提出的霍普菲尔德网络^[4]。理论上,递归神经网络模型可以获取到时间点为止的所有信息,并保存为隐藏状态。然而,在实际应用中,普通的递归神经网络模型并没有这么好的效果,尤其是对于比较长的输入,很难学习长距离依赖特征^[5]。RNN 模型如下图所示。

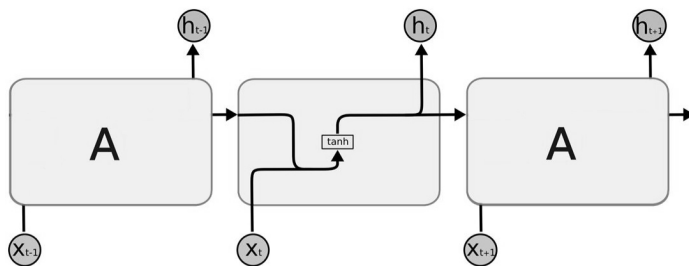


图 1 RNN 模型

2.2.2 LSTM 模型

长短期记忆网络通常只称为 LSTM,是一种特殊的 RNN,它能够学习长期的规律。它们是由 Hochreiter & Schmidhuber (1997) 首先提出的^[6],并且在后来的工作中被许多人使用和推广。LSTM 明确旨在避免长期依赖性的问题。长时间记住信息实际上是他们的默认行为,而不是他们难以学习的东西。所有递归神经网络都具有神经网络重复模块链的形式。在标准 RNN 中,该重复模块将具有非常简单的结构,例如单个 tanh 层^[7]。LSTM 也具有这种类似链的结构,但重复模块具有不同的结构。有四个,而不是一个神经网络。LSTM 模型如下图 2 所示。

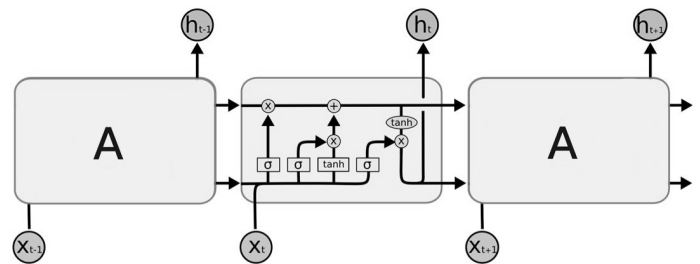


图 2 LSTM 模型结构

图中各种元素图标的解释如下图 3。



图 3 图标解释

3 实验与结果分析

3.1 模型选择

基于 DREAM 的商品推荐模型曾被广泛研究,该模型是把用户购物车中的商品向量化后构建购物车向量,将该购物车向量应用于神经网络模型继而得到用户购买特点的个性化向量。最终结果是通过商品向量和用户的个性化向量积来计算出商品购买概率。明显可以看出,此模型只考虑独立商品和用户之间的购买关系,但是并没有考虑用户的购买时序,也就是说没有考虑用户的购买历史。所以本文的商品推荐模型是基于 LSTM 模型,能够在模型训练中充分考虑用户时序信息。模型的特点如下:

3.1.1 商品重构模型是根据购物车数据得到的向量,在特征提取上是基于用 USERID 和商品 SKUID 为粒度,可以看出用户和商品是一对多的关系,得到多条特征样本。

3.1.2 矩阵不再是只有一个向量 ,USERID 和商品 SKUID 组成的二元组将作为一部分时间步长 时间步长的输入具有连续性 ,下一个时间步长基于上一个时间步长 ,这一点是商品重构推荐模型的重点。基于此种时间步长的单神经元结构预测得到用户是否会再次购买曾经已经购买过的购物车中的某个 SKUID 商品的概率。

3.2 数据预处理

数据预处理是将用户购买记录的源数据转换成能够用于模型训练的用户商品向量。One- Hot 编码是一种常用的数据预处理方式 ,本文也将采用这种方式 ,将商品映射为整数值 ,即每种商品唯一的商品 id ,然后再将商品 id 进行二进制向量表示。例如 {0,0,3,0,0...0}表示商品 ID 为 3 的商品。

4 实验与分析

4.1 实验环境

本章在深度学习框架 Keras 上实现了设计的实验 ,Keras 是一个开源的深度学习框架 ,对比较流行的深度学习框架 Tensorflow、Theano 等进行了高层 API 的封装 ,使用起来更加简便 ,使用的时候可以自由选择后端框架 ,本文实验所使用的具体实验环境如下表 1。

表 1 实验环境

操作系统	Ubuntu 14. 04
开发平台	Keras 深度学习框架
开发语言	Python
CPU	8 核 2. 10GHz * 2
硬盘	1TB
GPU	K80 * 2
内存	128GB

本文选择 Tensorflow 框架作为 Keras 的后端。相比于 Keras 框架 ,Tensorflow 框架的实验代码编写更加灵活 ,可以自由的定义模型结构以及损失函数。

4.2 实验数据与实验参数

实验数据集来于某新零售公司实际销售数据 ,时间从 2019 年 10 月到 2019 年 12 月 ,记录总条数 939060 条。其中用于模型训练的数据占 80% ,用于参数调优的数据占 10% ,用于验证结果的数据则是剩下的 10%。

4.3 评估指标

算法的评价指标是准确率(Precision) ,是指在推荐给用户的所有商品中用户实际购买的的商品所占的比例 ,见下式(1) :

$$Precision = \frac{\sum_{u \in U} R(u) \cap T(u)}{\sum_{u \in U} R(u)} \tag{1}$$

T(u)表示用户 u 实际购买的所有商品 ,R(u)表示推荐给用户 u 的所有商品。准确率越高 ,表示推荐的商品越符合用户的购买意向。

4.4 对比实验和结果分析

下面展示本文算法的训练结果 ,与 RNN 神经网络模型做一

个比较 ,以展示二者的差异和本文算法的优点 结果如下表 2。

表 2 推荐结果准确率比较

商品类别	RNN 推荐算法	本文算法	提高幅度 (%)
全部商品	0.047	1. 136	184. 8
1	0. 156	0. 169	8. 3
2	0.088	0. 211	139. 8
3	0. 249	0. 257	3. 2
4	0. 246	0. 277	12. 6
5	0. 271	0. 275	1. 5
6	0. 550	0. 580	5. 5

实验结果显示 ,在商品类别相近的商品推荐上本文算法效果更好 ,反之 ,效果近似。此外 ,在序列长度越长的训练场景 本文效果更好 ,因为 LSTM 能够考虑时序因素 ,将上一步长的结果作为下一步长的输入。

5 结论

本文根据用户在新零售电商平台重构商品行为和 LTMS 的特点 ,将 LSTM 模型应用于新零售商品推荐。

LSTM 神经网络是对 RNN 神经网络的优化 ,上一步时间步长作为下一步时间步长的基础 ,因此对于商品重构模型这种时序性要求较强的应用场景具有很好的训练效果。但是由于加入了遗忘门的存在 ,无法考虑全部历史信息 ,在某种检索场景 ,这是一个缺憾 ,一定程度上会对商品重构准确率造成影响。

在未来的研究中 ,需要不断调优参数 ,了解更多种类可以应用于商品推荐的模型 ,力求找到一种更好的模型 ,能够具有更高的商品重构推荐准确率。

参考文献

[1]黄立威 ,江碧涛 ,吕守业 ,等.基于深度学习的推荐系统研究综述[J].计算机学报 ,2018 ,41(7):1619-1647.
[2]梁昌勇 ,冷亚军 ,王勇胜 ,等.电子商务推荐系统中群体用户推荐问题研究[J].中国管理科学 ,2013 ,21(3):153-15.
[3]李浩君 ,张广 ,王万良 ,等.基于多维特征差异的个性化推荐方法[J].系统工程理论与实践 ,2017 ,37(11):2995-3005.
[4]陈卓 ,杨炳儒 ,宋威 ,等.序列模式挖掘综述[J].计算机应用研究 ,2008 ,25(7):1960-1963.
[5]刘吉.小世界特征的卷积神经网络模型研究与应用[J].2018,40(6):1517-1567.
[6]王栋,李业刚,张晓.基于多神经网络协同训练的命名实体识别[J].智能计算机与应用刊,2020,20(2):143-14.
[7]中国互联网信息中心.(2019).中国互联网发展状况统计报告.[EB/OL].Available at:http://www.cnnic.net.cn/hlwfzyj/hlwxbzg.