

万家山. 基于深度学习的混合主题模型应用[J]. 湖南科技大学学报(自然科学版), 2020, 35(3):102-109. doi:10.13582/j.cnki.1672-9102.2020.03.015

Wan J S. Application of Hybrid-Topic Model Based on Deep Learning [J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2020, 35(3):102-109. doi:10.13582/j.cnki.1672-9102.2020.03.015

基于深度学习的混合主题模型应用

万家山*

(安徽信息工程学院 大数据与人工智能学院,安徽 芜湖 241000)

摘要:在线社交网络的快速发展得益于主题模型的广泛应用.然而,目前几种典型的主题模型存在需要手动调参、语义连贯性不足、特征提取不充分和样本效率低等问题.对此,构建主题模型时利用深度学习技术来进行主题划分,以 BiLSTM-CNN 模型框架为基础,并在主题特征提取进行主题划分阶段引入“作者-主题”模型进一步优化框架,从而提出了一种混合主题模型 Hybrid-Topic Model (简称 HTM).将 LDA, CNN, BiLSTM-CNN 和 HTM 这 4 种主题模型应用于 2 组不同场景的数据集,并对结果进行对比分析.分析表明,在主题分类效果和 content 困惑度方面,HTM 主题模型的效果明显优于现有模型.此外,该模型在样本使用效率和模型迁移学习能力方面也有出色的表现,为后期研究指明了方向.

关键词:主题模型;LDA;角色识别;深度学习;社区推荐;HTM

中图分类号:TP311 文献标志码:A 文章编号:1672-9102(2020)03-0102-08

Application of Hybrid-Topic Model Based on Deep Learning

Wan Jiashan

(School of Big Data and Artificial Intelligence, Anhui Institute of Information Technology, Wuhu 241000, China)

Abstract: The rapid development of online social networks was attributed to the extensive application of topic models. Some typical topic models still had such problems as the manual regulation of parameters, insufficient semantic coherence, feature extraction deficiency and low sampling efficiency. For that reason, the deep learning technique was applied to topic division. The Author-Topic model was applied to topic feature extraction to optimize the BiLSTM-CNN model framework, a Hybrid-Topic Model (HTM) was proposed. LDA, CNN, BiLSTM-CNN and HTM topic models were applied to two data sets of different scenarios and the results were compared and analyzed. The results indicated that in terms of topic classification effect and content confusion, HTM is significantly more effective than the existing models. In addition, the model provides excellent performance in terms of sample use efficiency and model transfer learning ability, which points out the direction for later research.

Keywords: topic models; LDA; roles identification; deep learning; social network recommendation; HTM

随着移动互联网的快速发展和智能手机的广泛使用,广大用户已经从信息匮乏的时代迈入了信息过载的自媒体时代.用户在社交网络上交流互动产生了海量数据,如何从海量信息中快速获取用户感兴趣的

收稿日期:2019-01-10

基金项目:国家重点研发计划课题资助(2017YFD0301303);安徽省教育厅自然科学重大项目资助(KJ2017ZD53);安徽省教育厅自然科学重点项目资助(KJ2017A799)

*通信作者, E-mail: jswan@iflytek.com

信息,推荐系统成了解决这一问题的关键.推荐系统不仅应用到社交网络领域,还应用到电子商务、音乐、电影和视频、基于位置服务和广告等领域,让人印象深刻的包括社交网络上的热门话题推荐、交友推荐和电子商务各类商品的推荐等.目前使用的推荐算法(见表1)主要有(1)基于内容的推荐^[1]; (2)协同过滤推荐^[2]; (3)基于关联规则的推荐^[3]; (4)基于知识的推荐^[4]; (5)混合推荐^[5].

表1 个性化推荐的应用实例

应用领域	实例	推荐算法
电子商务	Amazon, eBay, 淘宝, 京东	协同过滤、基于关联规则
网页标签	Fab, Foxtrot	协同过滤
电影	MovieLens, Netflix, 豆瓣, 爱奇艺, 优酷	协同过滤
音乐	Pandora, Ringo, 网易云音乐, 酷狗	基于内容、混合推荐
旅游	携程, 途牛	基于内容、基于知识
求职	卡斯帕尔 Casper, 智联招聘, BOSS 直聘	基于内容
社交	Facebook, Twitter, 今日头条, 知乎, 微信	协同过滤、基于知识

推荐系统虽然广泛地应用在各个领域,但随着数据规模的增加,普遍面临着矩阵稀疏性、推荐系统的脆弱性等问题,主要表现在冷启动、新物品、新用户的推荐和大数据处理与增量计算上.研究主要围绕推荐系统在社交网络中的应用重点展开.

基于社交网络的个性化社区推荐系统是在个性化推荐系统和复杂社交网络分析的基础上发展起来的,目前在线社交网络的快速发展积累了大量用户行为数据,但也普遍面临着数据稀疏性、冷启动与可扩展性等问题.目前,主要采用主题分类与划分的方式能够有效缓解冷启动问题,提升整个推荐系统精准性.如:今日头条在首次使用是根据热门话题和位置进行推荐,使用过程中机器识别出你感兴趣的主题或关键词,根据文本内容分类并进行推荐;知乎核心功能是围绕相关话题展开智能问答.

通过上述2个社区推荐系统可以看出主题模型得到广泛应用,这也让主题模型成为近几年研究的热点.主题模型是文本语义信息的抽取模型,也是文本语义表征的有效方法.通过主题建模,不仅可以识别出隐含在文本中的主题语义单元,还能够将文本表示在语义信息更为丰富的主题特征空间上,从而有助于文本分类聚类、突发事件检测、主题演化分析、推荐系统等任务.

1 相关工作

1.1 几种典型的主题模型

本文将阐述几种主流的主题模型分别有 LSA, pLSA, LDA, HDP, Word2vec 以及基于深度学习的主题模型(见表2),下面将围绕主题建模及其相关技术展开探讨.

表2 几种典型的主题模型优缺点

主题建模技术	优点	缺点
潜在语义分析 Latent Semantic Analysis (简称 LSA) ^[6-7]	1)方法快速且高效.能够快速计算出术语与文档的相似度.	1)缺乏解释的嵌入,且矩阵的秩要根据经验确定; 2)需要大量的文件和词汇来获得准确的结果; 3)表征效率低.
概率潜在语义分析 Probabilistic Latent Semantic Analysis (简称 pLSA) ^[8-9]	1)模型可解释性强,因定义了概率模型,使概率分布有明确的解释; 2)模型更加灵活,表征效果较好.	1)给新文档分配概率存在很大难度; 2)参数量随着文档数线性增长,容易出现过度拟合问题.
潜在狄利克雷分布 Latent Dirichlet allocation (简称 LDA) ^[10-12]	1)LDA 比 pLSA 效果更好,模型可以轻而易举地泛化到新文档中; 2)是目前使用最普遍的主题建模技术.	1)模型训练需要预先设定主题数目,根据结果手动调参,而且计算量比较大,训练时间比较长; 2)LDA 对短文本的主题分类效果较差.
层次狄利克雷过程 Hierarchical Dirichlet Process(简称 HDP) ^[13-15]	1)具有较好的鲁棒性和灵活性,能够自动确定主题数目.	1)存在部分主题会丢失的可能性; 2)忽略了时间信息对主题分类的影响.

续表 2

主题建模技术	优点	缺点
Word2vec ^[16-17]	1)模型泛化能力和可解释性强,模型具有强大的表征能力; 2)训练速度快,模型不需要标注数据,通过训练能够学习出词与词之间的相似性和内在联系.	1)模型受语料来源的影响较大; 2)需要大量的训练数据来调整优化模型并且需要避免过拟合,训练模型时容易引发维度灾难.
深度学习主题模型(TextRNN, TextCNN, LSTM等) ^[18-20]	1)模型具有强大的学习能力,模型泛化能力强; 2)表征效果高.能够将文档向量和词汇向量协同起来,不仅能学习词汇的词嵌入,还同时学习主题表征和文档表征,并反映出两者间的关系.	1)大部分深度学习主题模型对样本数据集的大小要求,需要大量的训练样本才能获得较好的分类效果; 2)由于深度学习的需要,模型对硬件具有相对较高的要求.

1.2 LDA 主题模型

目前,具有代表性的主题模型隐含狄利克雷分布(Latent Dirichlet Allocation, LDA)是由 Blei^[10]于2003年提出的,该模型利用无监督的学习方法挖掘文本中隐含的主题信息,并将文档集中每篇文档的主题以概率分布的形式给出,从而根据主题分布进行主题聚类或文本分类,这种方法优势在于不需要任何关于文本的背景知识,直接可以根据主题分布进行主题聚类或文本分类.从生成模型的角度来看,每个文档中的每个词汇都可以表示成“文档按照一定概率来确定属于某个主题,并根据这个主题以一定概率选择某个词汇”这样一个过程.那么通过逆向考虑,一篇文档中每个词汇出现的概率可以表示为

$$p(\text{词汇}|\text{文档}) = \sum_{\text{主题}} P(\text{词汇}|\text{主题}) \times P(\text{主题}|\text{文档}). \tag{1}$$

为了更加直观地看出文档、主题以及词汇三者之间的关系,现将 LDA 模型具体表示为图 1.

从图 1 可以看出将 LDA 三层结构(图中符号的含义见表 3), z_n 由 θ 生成,则 $P(z_n|\theta)$ 表示每个文档所产生主题 z 的概率分布情况. $w_{m,n}$ 由 $z_{m,n}$ 和 β 共同生成,则 $P(w_n|z_n, \beta)$ 为每个主题所对应词汇 w 的概率分布情况.通过上面对 LDA 生成模型的讨论,可以知道 LDA 模型主要根据学习训练调节控制参数 α 和 β 并最终确定模型.

$\alpha \rightarrow \theta_m \rightarrow \text{No}_{m,n}$ 这个过程表示在生成第 m 篇文档的时候,先抽取一个主题 θ_m , 然后根据文档获得第 n 个词的主题编号 $\text{No}_{m,n}$.

$\beta \rightarrow \varphi_k \rightarrow \text{No}_{m,n} \rightarrow w_{m,n}$ 这个过程表示在主题 φ_k 中,挑选编号为 $k = \text{No}_{m,n}$ 的主题进行查找,然后生成词汇 $w_{m,n}$.

$$p(\theta, z, w|\alpha, \beta) = p(\theta|\alpha) \prod_{n=1}^N p(z_n|\theta) p(w_n|z_n, \beta). \tag{2}$$

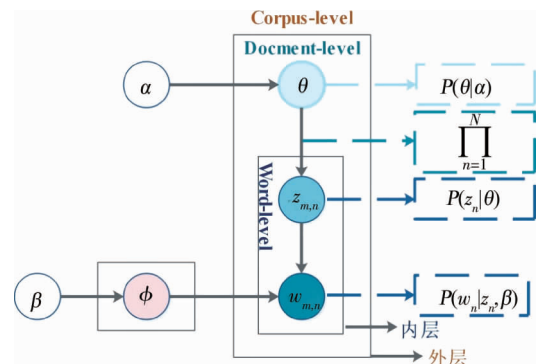


图 1 LDA 模型

表 3 符号含义表

符号	含义
k	主题编号
m	文档编号
n	词汇编号
α	狄利克雷分布的参数,表示文档-主题分布
β	先验参数,表示主题-词汇分布
$z_{m,n}$	第 m 个文档中第 n 个词汇的主题
$w_{m,n}$	第 m 个文档中第 n 个词汇
$\text{No}_{m,n}$	第 m 个文档中第 n 个词汇的主题的编号
φ_k	第 k 个主题下词汇的分布
θ_m	第 m 个文档下主题分布

2 构建一种混合主题模型 HTM

2.1 角色识别

用户更喜欢来自朋友的推荐而不是被系统算出来的推荐,社会影响力被认为比历史行为的相似性更加重要^[21]。“美国著名的第三方调查机构尼尔森调查了影响用户相信某个推荐的因素.调查结果显示,90%的用户相信朋友对他们的推荐,70%的用户相信网上其他用户对广告商品的评论。”^[22]可以看出,好友或具有经验的用户(定义为“被信任者”或“领袖角色”)的推荐对于用户(信任者)是否采纳推荐结果有着重大的影响.每个用户在社会网络中都承担着各自的社会角色,那么如何针对用户进行角色识别(即为社会网络中的社会角色识别)就显得至关重要.

在研究 LDA 方法的基础上,针对用户进行角色识别,通过“文档-作者”矩阵可以确定不同主题的综合权重,并将这些方法进一步扩展到作者建模的过程中,识别出作者及与作者相关的主题.这样在原有的“文档-词汇”“主题-词汇”“文档-主题”3个矩阵中增加“作者-主题”矩阵.通过引入“作者-主题”模型搭建了“用户”和“主题”的桥梁,让主题模型能够识别出用户之间的相关性,并推算出以主题为中心的重要用户(即潜在领袖角色)和用户群.

如图 2b 所示,其中 x 表示具体给出指定词汇的作者.“作者-主题”生成模型(Author-Topic Model)的主要思路:首先,从作者集合 A 中随机选择,每一个作者对相应指定词汇的概率分布表示为 φ ;其次,根据估计 φ 的值和作者相应的权重信息,得到主题 z 和 φ 分布对应主题的词汇 w ,并获取关于作者的感兴趣信息;最后,借鉴协同过滤算法思想得到相似兴趣的作者信息以及与该文档主题具有相似性的作者群组.

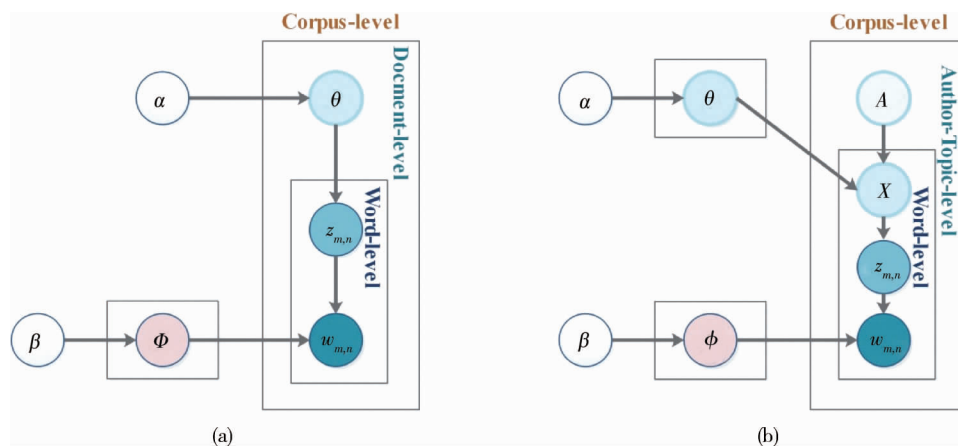


图2 “作者-主题”生成模型

“作者-主题”生成模型融合了 LDA 主题模型和作者模型的主要思想:通常“每个文档都有一个唯一的作者相对应,每个作者都有一个唯一的主题相对应”.据此,可以获取作者通常关注或擅长的主题信息,并推送这些主题所代表的相关文档.

2.2 混合主题模型 HTM

传统的概率主题模型由于其浅层的特征结构以及概率化的生成模式依旧面临着模型可扩展性、主题语义连贯性、推断上下文一致性、特征表达能力等不足.深度学习技术的不断成熟为自然语言处理领域带来了新的发展契机,也为主题模型提供了新的构建思路.长期以来,词向量嵌入一直都是自然语言处理的核心表征技术,预先训练词向量给自然语言处理领域的发展取得了显著成效,并用于机器翻译、问答系统和主题提取等广泛的 NLP 任务.此外,语言建模、神经网络等深度学习方法在文本语义特征表达方面取得了突破性进展,为构建深层次且具有语义连贯性的主题模型提供了可能.

依托深度学习框架,通过 CNN 训练词向量^[23]来提取句子中的主干信息为文本分类奠定基础.一般而言,CNN 具有捕获词级别的 n-gram 能力,用在句子分类任务中效果奇佳,LSTM 用来处理具有时序依赖特征的序列中也有奇效.然而,当 LSTM 用在超长类型的文本文档上,对于梯度弥散问题也是无力回天.针对

梯度爆炸的情况,常常会使用截断梯度方法.但是梯度截断并不能有效地处理梯度消散问题,目前常见的方法是使用正则化或约束参数^[24].

随着深度学习的深度应用,对数据量的要求在不断加大,在无法获得大量数据的背景下,迁移学习能力和样本使用效率成了制约模型应用的重要因素.把主题识别看成是多分类问题,为了提高样本使用效率,通过微调预训练模型,选择表达性强的语言模型编码器(如深度 BiLSTM 或 Transformer),提升模型的迁移学习能力.

本文为了发挥 CNN 和 BiLSTM 模型的各自优势,使用 Bi-directional LSTM 模型,不仅可以很好地表征出整个文档,而且能够提高模型文档分类效果和迁移学习能力.同时,考虑到用户角色之间的影响,在主题特征提取进行主题划分阶段引入“作者-主题”模型,提取用户感兴趣的主题进一步优化 Bi-directional LSTM 主题模型形成混合主题模型(简称 HTM),HTM 模型框架如图 3.

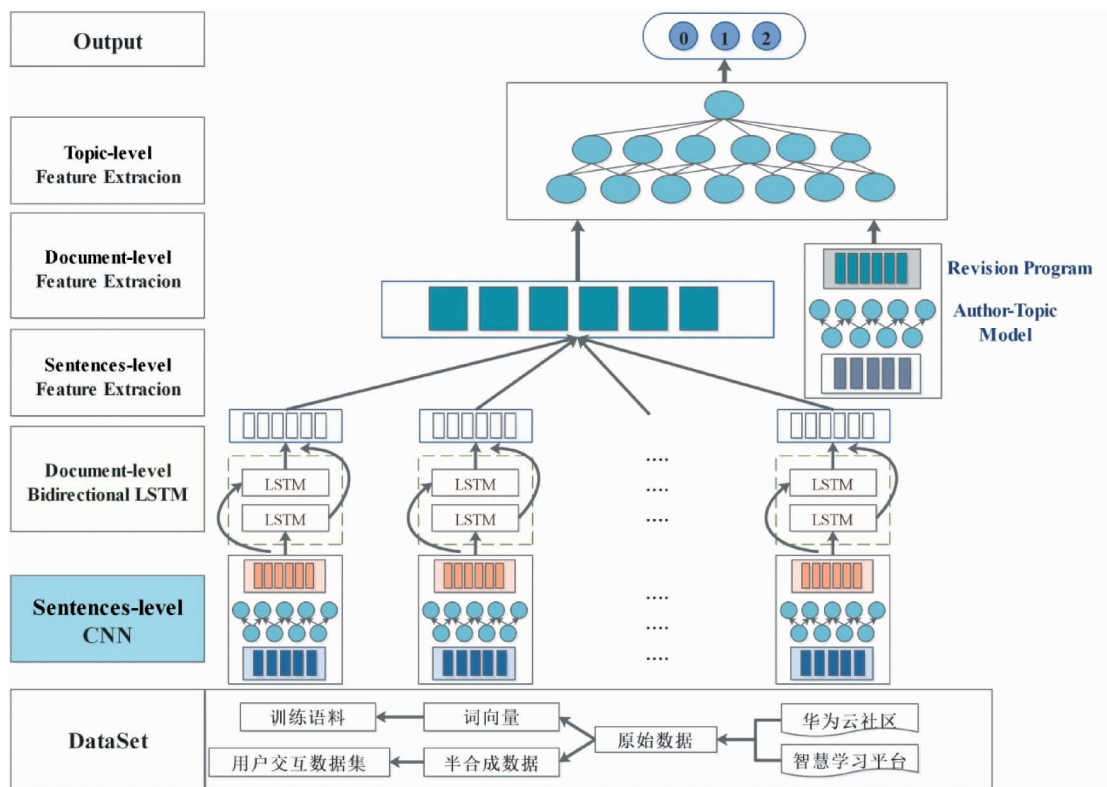


图 3 混合主题模型框架

该混合主题模型可以同时学习主题表征、文档表征和作者信息表征,具有较强的特征提取和资源表征能力,可以大幅度提高样本使用效率,只需要较少的样本数据就可以达到最优性能.因此,该模型在训练示例数量少、注释成本高的领域具有较好的应用前景.

3 数据集与评估方法

本文使用的 2 组原始数据分别来自华为云社区 (huaweicloud.com) 和智慧学习平台 (iflysse.com) 社区问答数据,用于主题模型的验证和对比.

3.1 数据集及模型参数标定

实验数据集包含了华为云社区 2017 年 12 月 7 日利用 Scrapy 抓取的 3 500 个网页和智慧学习平台 2018 年 8 月 18 日导出的 1 300 条处理后的问答数据.选取 LDA, CNN, BiLSTM-CNN, HTM 主题分类模型进行对比,LDA 的主题数目 K ,狄利克雷分布参数 α 和 β ,经过实验训练对相关参数调优并分别设置为 50, 0.5 和 0.02.CNN 模型隐含层数 1 层、隐含层神经元个数 600 个、卷积核为 3、dropout rate = 0.5.

3.2 评估指标

3.2.1 分类错误率.

当文档按照主题对应具体数值来分类输出时,通常评估分类的指标有准确率或错误率指标来计算预测值和真实值之间的误差,此处采用分类错误率指标(Test error rates),具体定义公式为

$$\text{Test error rates} = (1 - \text{正确分类文档数} / \text{预测分类文档总数}) \times 100\% .$$

3.2.1 内容困惑度.

内容困惑度指标被经常应用于主题模型的效果评价中^[14],主题模型困惑度越低,则表示主题模型效果越好.其中, $p(w)$ 为每一个词汇出现的概率, N 为词汇的总数,得到计算公式为

$$\text{perplexity} = \exp \frac{-\sum \log(p(w))}{N} . \quad (3)$$

4 实验与结果分析

介绍实验的操作过程,对比不同推荐模型和方法在2组数据集上的评估指标变化情况,并讨论如何结合主题找到相关用户提升分类效果.

4.1 实验相关设置

数据预处理与划分:本文数据集包含2组数据,将数据集划分为训练集和测试集,其中训练集占整个数据集比例的80%.进行了10次的随机划分后,用10次测试结果的平均值来评估模型的预测质量.

实验软硬件平台搭建: Anaconda 5.2.0, TensorFlow 1.5.0rc1, PyCharm Version 2017.3.4, MATLAB R2012a, Inter(R) Core(TM) i7-8550U CPU@1.80 GHz, 内存(RAM)16.0 G.

4.2 模型对比分析

LDA 由于未考虑云社区数据的特点,抽取的主题可读性不强,不利于用户理解.另外,LDA 需要预先设置主题的数目,这样人为因素的影响较大,与实际主题分布的情况相符的概率较小.

CNN 一般利用词向量初始化模型,该模型需要从头开始学习,不仅要学习消除歧义,还要学习从词组成的句子中提取意义,模型需要大量的示例才能获得良好的性能.对于应用在训练示例数量少、注释成本高的领域就很难施展.

BiLSTM-CNN 由于使用了 BiLSTM 提升模型的迁移学习能力,使得训练模型需要的数据比从头训练模型的数据要少.但 BiLSTM 因为模型容量问题,在学习较长句子时,可能丢弃一些重要信息导致特征提取不够充分.

HTM 混合主题模型是在 BiLSTM-CNN 模型基础上,考虑到用户之间的相关影响,加入“作者-主题”模型进行用户角色识别优化 BiLSTM-CNN 模型.该混合主题模型不仅具有较好的迁移学习能力,而且具有较强的特征提取和资源表征能力,可以大幅度提高样本使用效率,从而实现较少的样本数据就可以达到良好的性能.

4.3 实验结果与分析

实验 1. 分类错误率(全体样本)

选取全部样本数据集,将 CNN, BiLSTM-CNN 和 HTM 模型对比传统的主题分类 LDA 模型的结果如表 4 所示,实验结果表明传统的主题分类模型误差大于深度学习的主题模型.对比 2 组数据可以看出同时采用传统主题分类模型,对原始数据样本存在一定的依赖性,在华为云社区和智慧学习平台 2 组数据集对比误差存在着较大的差距,主要原因在于 LDA 主题模型对短文本的主题分类效果不好.

表 4 4 种主题模型在 2 组数据集集中的分类效果

Model		Test error rates/%	Model		Test error rates/%
Huaweicloud	LDA	9.3	Iflysse	LDA	15.6
	CNN	6.0		CNN	8.9
	BiLSTM-CNN	5.9		BiLSTM-CNN	8.6
	HTM	5.4		HTM	6.3

实验 2. 分类错误率(抽取不同数量样本)

选取部分样本数据集,对比 4 种主题分类模型的结果见图 4 所示,实验结果表明 4 种主题模型都受到训练样本的数量变化而变化,其中可以很明显地给这 4 种主题模型按照变化从大到小排序,分别是 $LDA > CNN > BiLSTM-CNN > HTM$.从数据集特点来看,由于 b 数据集相对于 a 来说,单个文本的篇幅较小,且讨论的主题较泛.对比图 4a 和图 4b 可以得出,LDA 模型受数据集特点影响较为明显,HTM 模型受数据集特点影响较小,能够在样本较小的情形下取得相对令人满意的分类效果,该模型稳定性最强,能够有效地缓解因矩阵稀疏性引起的冷启动问题.

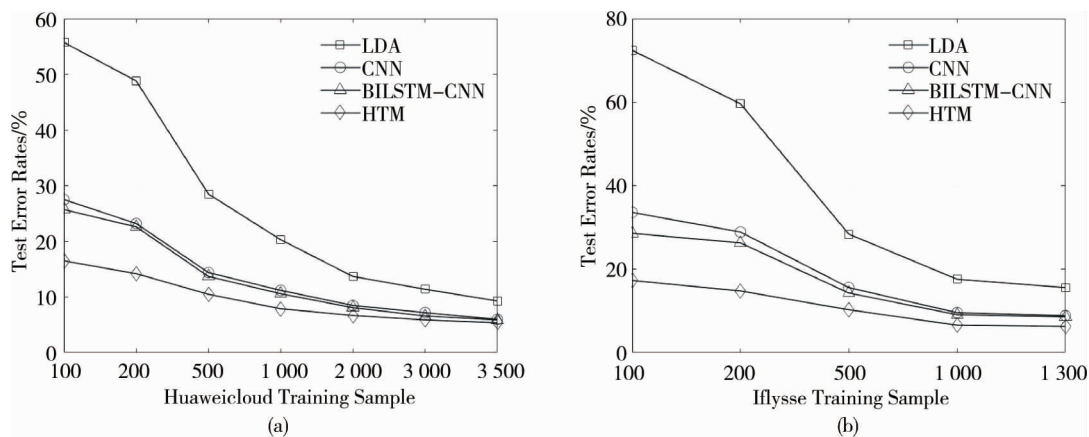


图 4 4 种主题模型分类效果

因考虑实验效果的可验证性,针对实验 3 采用智慧学习平台的数据集进行分析.

实验 3. 主题数目与内容困惑度-iflysse 样本

采用智慧学习平台数据集,由图 5 看出主题分类一旦超过 60 时,随着主题挖掘程度的不断提升,各个模型的困惑度变化不明显.从整体趋势看 HTM 主题模型的困惑度最低,得到的主题模型效果越好.

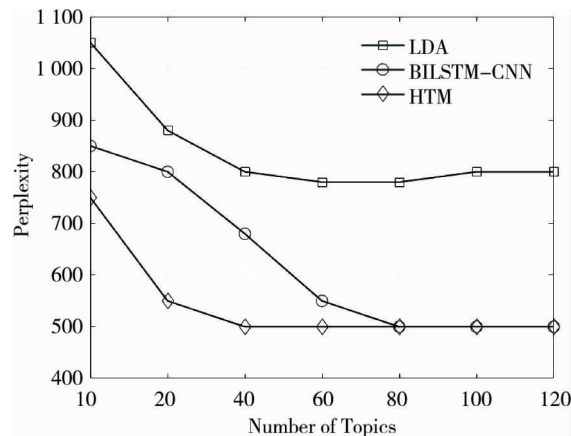


图 5 模型的内容困惑度

5 结论

- 1) 根据在 2 组数据集下的应用对比,分析发现 HTM 模型需要的数据不仅比从头训练模型的数据要少得多,而且错误率低、模型效果好.
- 2) 通过引入“作者-主题”模型,使得 HTM 模型能够在长短不同的文本中表现出更好的稳定性.
- 3) 利用 BiLSTM-CNN 和“作者-主题”优化模型,使得模型在内容困惑度方面表现出更快的收敛性,从而达到更为出色的预测效果.

最终的实验结果表明,文中的几处改进都提升了模型的应用效果.但本文并未考虑到改进后的模型复

杂度,后续工作将侧重模型范式进行优化,在不影响现有应用效果的基础上,寻求更加简洁、高效的模型框架。

参考文献:

- [1] Pazzani M J. A framework for collaborative, content-based and demographic filtering[J]. *Artificial Intelligence Review*, 1999, 13(5/6): 393-408.
- [2] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborative filtering recommendation algorithms[C]// *International Conference on World Wide Web*. ACM, 2001: 285-295.
- [3] Mican D, Tomai N. Association-rules-based recommender system for personalization in adaptive web-based applications [C]// *International Conference on Web Engineering*, Springer, Berlin, Heidelberg, 2010: 85-90.
- [4] Blanco-Fernández Y, Pazos-Arias J J, Gil-Solla A, et al. A flexible semantic inference methodology to reason about user preferences in knowledge-based recommender systems[J]. *Knowledge-Based Systems*, 2008, 21(4): 305-320.
- [5] Thorat P B, Goudar R M, Barve S. Survey on collaborative filtering, content-based filtering and hybrid recommendation system [J]. *International Journal of Computer Applications*, 2015, 110(4): 31-36.
- [6] Foltz P W. Latent semantic analysis for text-based research[J]. *Behavior Research Methods Instruments & Computers*, 1996, 28(2): 197-202.
- [7] Dumais S T. Latent semantic analysis [J]. *Annual Review of Information Science & Technology*, 2010, 38(1): 188-230.
- [8] Brants T, Chen F, Tsochantaridis I. Topic-based document segmentation with probabilistic latent semantic analysis[C]// *Eleventh International Conference on Information and Knowledge Management*. ACM, 2002: 211-218.
- [9] Hofmann T. Unsupervised learning by probabilistic latent semantic analysis[J]. *Machine Learning*, 2001, 42(1/2): 177-196.
- [10] Blei D M, Ng A Y, Jordan M I. Latent dirichlet allocation[J]. *Journal of Machine Learning Research*, 2003, 3: 993-1022.
- [11] 李天辰, 殷建平. 基于主题聚类的情感极性判别方法[J]. *计算机科学与探索*, 2016, 10(7): 989-994.
- [12] 何伟林, 谢红玲, 奉国和. 潜在狄利克雷分布模型研究综述[J]. *信息资源管理学报*, 2018(1): 55-64.
- [13] Teh Y W, Jordan M I, Beal M J, et al. Hierarchical dirichlet processes [J]. *Publications of the American Statistical Association*, 2006, 101(476): 1566-1581.
- [14] 刘少鹏, 印鉴, 欧阳佳, 等. 基于 MB-HDP 模型的微博主题挖掘[J]. *计算机学报*, 2015, (7): 1408-1419.
- [15] 李永忠, 郑滔. 基于标签的半监督 HDP 文本分类主题模型[J]. *模式识别与人工智能*, 2017(12): 1138-1148.
- [16] 刘耘. 基于词嵌入的模糊文本搜索与推荐系统[D]. 南京: 东南大学, 2017.
- [17] Wang Z, Ma L, Zhang Y. A hybrid document feature extraction method using latent Dirichlet allocation and Word2Vec[C]// *IEEE First International Conference on Data Science in Cyberspace*. IEEE Computer Society, 2016: 98-103.
- [18] 张闯. 基于深度学习的知乎标题的多标签文本分类[D]. 北京: 北京交通大学, 2018.
- [19] 李洋, 董红斌. 基于 CNN 和 BiLSTM 网络特征融合的文本情感分析[J]. *计算机应用*, 2018, 38(11): 3075-3080.
- [20] 朱艳辉, 李飞, 冀相冰, 等. 反馈式 K 近邻语义迁移学习的领域命名实体识别[J]. *智能系统学报*, 2019(4): 820-830.
- [21] 潘一腾, 何发智, 于海平. 一种基于信任关系隐含相似度的社会化推荐算法[J]. *计算机学报*, 2018(1): 65-81.
- [22] 项亮. *推荐系统实践*[M]. 北京: 人民邮电出版社, 2012: 144-164.
- [23] Xu H, Dong M, Zhu D, et al. Text classification with topic-based word embedding and convolutional neural networks[C]// *Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. ACM, 2016: 88-97.
- [24] Zhang L, Xiang F. Relation classification via BiLSTM-CNN[C]// *International Conference on Data Mining and Big Data*. Springer, Cham, 2018: 373-382.