

结题报告

——基于深度模型的影评情感分析

小组成员：

赵冬迪 2120161081 解亚东 2120161004

夏光敏 2120161064 杨雅婷 2120161070

内容与意义

随着网络上评论信息爆炸式的增长，以人工的方法很难应对海量评论信息的收集和处理，情感分析技术随之产生。本课题以电影影评为例，拟采用深度学习的方法对电影评论进行情感分析。

一个电影评论网站允许用户提交关于他们针对某部电影的带有感情色彩的主观评论。充分挖掘这些评论继而生成有价值的元数据，可以让我们以一个宏观的角度来理解观众对于这部电影的情感，我们可以对主观内容做出一个客观的分析，使我们能更好的理解产品和服务的趋势，为观众和其他电影参与方提供更好的决策建议。相比较于单纯的评分，基于影评内容的分析更能体现一个人的主观情绪。

通过训练得到的影评情感分析模型，可以对微博、Twitter 等社交媒体中包含的影评内容进行分析，从而理解不同类型用户对于电影的真实情感。

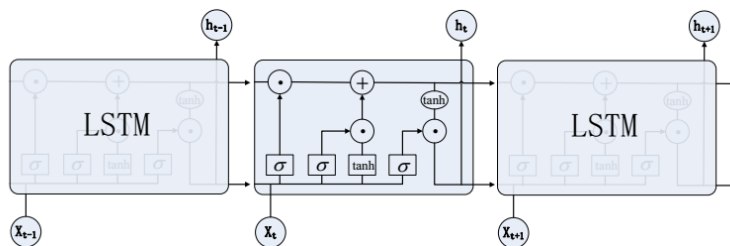
目标与功能

我们的目标是利用深度模型进行影评情感分析，训练出来的情感评估模型可以用作以下用途：

- 用户电影推荐
- 影院电影排片推荐
- 电影种类受欢迎程度评估

模型介绍

LSTM 网络是 RNN 网络的扩展，专门设计用来避免长期以来问题。LSTM 的循环神经网络模块具有不同的结构，这与传统的 RNN 网络不同，存在 4 个以特殊方式相互影响的神经网络层，网络模块示意图如下：



LSTM 网络的关键在于细胞状态。在 LSTM 中，通过门结构来对细胞状态增加或删除信息，而门结构是选择性让信息通过的方式，通常有一个 sigmoid 神经网络层和逐点乘积操作组成。

LSTM 网络具有输入门、忘记门和输出门等三种门结构，用以保持和更新细胞状态。LSTM 网络模型已被成功地应用到诸如文本/情感分类、机器翻译、智能问答和看图说话等自然语言处理任务中。

实验过程与结果展示

1. 数组准备

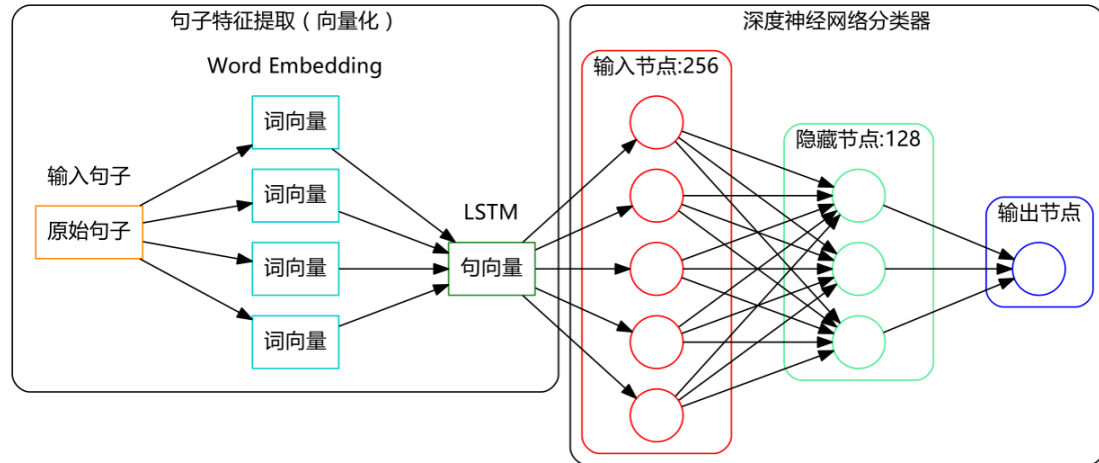
利用爬虫程序从豆瓣等影评网站爬取数据，并从中抽取用户名、评分、影评等内容。训练情感分类模型，并基于这些内容挖掘相关知识。

目前已经按照电影类别从豆瓣上爬取了近 20 万条影评作为模型的训练与测试集，主要涵盖了科幻、爱情、动画、文艺、战争、恐怖、纪录片、烂片八大类，数据的清洗已经完成，将标签不明显、不完整的数据删除。由于豆瓣的影评等级在 1-5 星之间，从爬取到的影评中抽取 4 星及以上的影评作为正面情感的样本，2 星及以下的影评作为负面情感的样本。处理后的数据集包含 8 万余条正样本和 5 万余条负样本。

1	电影名称	导演	演员	豆瓣评分	用户	评论星级	评论内容
2	美女与野兽	比尔康顿	主演 艾玛	7.4	斯文看客si	4	导演你把大表哥请来却不给他时间秀颜值，你这么做良心不会痛么？
3	美女与野兽	比尔康顿	主演 艾玛	7.4	晨下	4	当他受伤倒在雪地里时，我一下被萌翻了！因为他的脚爪有肉垫！！我的妈该肉垫！！毛绒绒！有尾巴！能单手托起妹子！最关键的是有爪爪有肉垫！！女巫你给！
4	美女与野兽	比尔康顿	主演 艾玛	7.4	王大根	3	Emma Watson一副街头小混混的样子搞得整个故事看起来完全不像是爱情片，更像是“我洪兴十三妹带刀来救我老豆，但后面感觉你野牛很讲义气所以我决定以后跟你混了。”
5	美女与野兽	比尔康顿	主演 艾玛	7.4	唱啦啦	3	谢天谢地后厨没演拉黑，全程画面单调干柴，另外迪士尼音乐真是有铁！
6	美女与野兽	比尔康顿	主演 艾玛	7.4	野野野野野	2	如坐针毡两小时，后厨小姐什么时候开始只唱得演歌，鼻尖到眉毛一动不动的呃？
7	美女与野兽	比尔康顿	主演 艾玛	7.4	马庆云	4	我跟野兽之间，只差着一个碧桂园的森林别墅。
8	美女与野兽	比尔康顿	主演 艾玛	7.4	冰川公爵	4	童话电影要什么深度内涵，观众就想看着梦幻公主和帅气的王子在高大的城堡里谈情说爱，艾玛美大表哥，那件惊艳的黄色蓬松公主裙更是！
9	美女与野兽	比尔康顿	主演 艾玛	7.4	凌睿	4	若你喜欢怪人，其实我很美。
10	美女与野兽	比尔康顿	主演 艾玛	7.4	羚羊的灵感	4	被加斯顿迷得神魂颠倒我还有救么---ps 因为爱乐之城黑屋后厨的人能再low一点吗？另外说改编不多就给差评的，不妨去看14年的原版啊，希望你们喜欢=)
11	美女与野兽	比尔康顿	主演 艾玛	7.4	撒特快	2	看得真犯尴尬感，这一直真是非常平庸的翻拍，而且或许是动画版实在太经典了，导致迪士尼在这次改编上几乎是束手束脚毫无作为的，对动画版的致敬确实还凑合。
12	美女与野兽	比尔康顿	主演 艾玛	7.4	蓝梓建就拜	4	一群英国人演了一群法国人，居然多了四首歌，而且这些唱得真是最终极大招啊。
13	美女与野兽	比尔康顿	主演 艾玛	7.4	MrMoonlit	5	为什么英美演员唱歌都这么好？连Emma Thompson阿婆和万磁王老先生都能唱，后厨小姐也不错，然后都唱成这样了还能守阵地不转出影视歌三栖巨星？让大表哥唱！
14	美女与野兽	比尔康顿	主演 艾玛	7.4	Ellie	4	画面超级美，艾玛的贝尔也是美出天际，电影没有太大惊喜，喜欢那个小茶杯，可怜我大表哥的盛世美颜只出场了2分钟，但是通过野兽看到丹尼尔的眼睛也是！
15	美女与野兽	比尔康顿	主演 艾玛	7.4	独立自主	4	大表哥最后出场不到5分钟，蓝色蝴蝶结让我不知如何是好。
16	美女与野兽	比尔康顿	主演 艾玛	7.4	lee	4	什么是绝世美颜？就是演了一整场野兽只用一个镜头就让电影院观众惊呼“啊”的大表哥。
17	美女与野兽	比尔康顿	主演 艾玛	7.4	too soon	5	野兽这么帅，为什么变成人类后贝尔毫不犹豫一秒就接受了？
18	美女与野兽	比尔康顿	主演 艾玛	7.4	tiger存存	5	这颜值美不行。
19	美女与野兽	比尔康顿	主演 艾玛	7.4	二月鸟语	2.4/10	导演太没创意了，一个很好的改编题材交到你手里，你却四平八稳的把动画还原成真人版，迪斯尼当初花那么多钱肯定不是让你这么敷衍了事的啊，剧情过于
20	美女与野兽	比尔康顿	主演 艾玛	7.4	伊谢尔伦	3	美国公司找两个英国人主演一个法国故事，闲聊的不是莎士比亚就是亚瑟王，我如果是法国观众我是我要投诉迪士尼的！！
21	美女与野兽	比尔康顿	主演 艾玛	7.4	kakakar!	4	真迪斯尼真人版，每年都要被迪士尼唤醒一少女心，比法国版合情合理很多，读书改变命运的宣传片，不读书的村民才是野兽，艾玛超美的，啧啧啧啧好美丽！

2. 模型设计与训练：

设计采用了基于 LSTM 神经网络的情感分析模型。模型的具体流程如下图：



由于爬取到的影评样本数据集是不均衡的，即正负样本的数量差距较大。因此，首先从数据集中随机抽取相同数量的正负样本生成训练数据集，再从剩下的数据集中抽取测试样本。

正负样本示例如下：

699	温馨结局看到和爸爸谈话的那段哭了	
700	无爱的达西 疯狂的爱情	
701	以为很平庸 结果意外的好看。	
702	让我想起了我和老公最后走到一起的感觉 实在很像 勾能很多回忆	
703	和Becoming Jane时一样的感动让我抵接欧歌，黄色舞鞋舞裙长清河逆流才送来一个Mr Darcy，是否有一次奥斯汀就象一次男主角，Anyway，你们的神色有相近所以让我觉得亲切，短剧要成功，	
704	相比95版我更喜欢这版，因为我太太太喜欢Keira Knightley这朵英伦玫瑰了，外表很美，布置很用心，摄影很漂亮，私心最喜欢Darcy在清晨追着朝阳去找Elizabeth的那段，阳光只爱人物的感情，刚刚好。	
705	我承认我是一个颜狗，《傲慢与偏见》是我看的第一本小说，也是至今都还认为不错的一本书，然而，之前被剧版的男女主角深深的伤害过，害得我都对这面有心理阴影了，这帮里面女的漂亮，男的不丑，衣服也不错，看得很开心。	
706	当两个无比骄傲的人终于发现自己的愚蠢，爱情才超越一切从天而降，祥和美好，片尾，Darcy一声接一声，一边索吻一边温柔地叫着Mrs Darcy，真是美到爆啊。	
707	真是喜欢Keira这个姑娘女子 新版很好看	
708	除了美见的下颚骨都很好的	
709	在我看来就是两个相爱的人之间的拉锯战，只有互相在乎对方才会作死，小打小闹怡情，不管中国人还是外国人都一个德行，在爱情方面总不会屈服，当然也许这就是爱情的本质	
710	这个世界上最好的动画片	
711	这是一只安静和善良的故事么。	
712	把所有的好东西都给eva	
713	乡下文艺宅小孩儿追求都市女精英的故事——以及——	
714	我看哭了。。。	
715	很久没看过这么可爱的动画片了，瓦力什么都不会说，除了那一声声执拗的：Eve！他们在星空画出美丽的弧线，他们无意间守候了一份真正的爱情。	
716	没有引述真是大遗憾了。	
717	再也没有比《瓦力》更美的动画片了！谈到Walt E的来源，导演安德鲁·斯坦顿否认了“ET”或“霹雳五号”的说法（尽管外形很像），他只承认自己创造Walt E受到了“小台灯”（Luzo）、双筒望远镜、R2-D2、迪士尼乐园“星空之旅”的机器人	
718	人类在这个片子里面是处于完全被嘲笑的位置。（其实应该说美国人？）片中的超大型连锁企业BNL在未来已经统一了所有的高科技领域，而人类懂得完全依赖电脑的帮助才能吃垮拉倒的胖子们，而且从小就接受电脑的教育要消费	
719	确实也都为尔喝彩了	
720	准备第二遍	
721	小瓦力，六人组	
722	地球上最后一袋收的精髓瓦力walle原话是巨婴座吧。。	

406 狗屁剧情，还重返18岁。因为倪妮才看的此片，但是看到董建华那面瘫脸就够了，特么的我居然还看完了——

407 倪妮！你已被列入黑名单

408 演技很浮夸，王大伟竟然有点帅，董建华完全没有CEO高富帅的气质，土的神渣，全赖高露，偶尔笑一下简直吓死人

409 剧情真心挺扯的？但是感觉力不错，倪妮马苏挺好的，画面也很美，马苏真心搞笑担当～但是王大伟什么鬼，董建华全程一个表情能不能请他退出电影圈??

410 倪妮演技真心不错，苏嘉宁演的，剧情还行。

411 看着剧本列表是下过功夫的，不过拍摄和表演上又掉脚本降几档。前面拍摄有点low，巧克力的引子一点都不酷，也没说服力。中后段18岁与28岁的对话展开后渐入佳境，年龄穿越的自我觉醒把握住了，梦想是要大于爱情的，倪妮!

412 你知道吧，这绝对是这个电影最高评价!

413 是二代的七安和二代的未成年，可算是女性题材电影成功和失败的对比案例，七安展示了如何升华和血原著并用对演员，未成年片如其名，整原著般的改编和失控的演员足以毁掉一切。

414 看到一个13岁小男孩演技独好又能聊又斯文的内心表白还真把我恶心到了——

415 适合以五倍速观看，一星给画作

416 小破情情死日常乐的

417 空洞，我知道导演想唤起我的什么感受，但是，那一部分感受我一定要通过观看的子和温才能被唤醒，动画这种形式，在这个目的前大失败了

418 怎么说呢，其实就是普通中学生浪漫爱情的那些事儿，没觉得怎么打动人，男主角一副欠扁的样子

419 这片是不是就一男两女？已经无聊到不想看的地步了&又看到很多用片名做标签的怪人

420 好废柴好废渣好废神

421 神经病片子 28片

422 太烂情

423 喜欢你谁懂啊

424 狗男女抽烟而已评分那么高有病啊

425 3.5星，从表演到情感都非常可爱，但格局还是破掉了点儿，感觉很有香港那种小尖酸，大梦想的城市气质。

426 说不是David Fincher的迷，许多时候他营造的所谓“剧情”在我看来不像是melodrama，通俗亦平庸，所谓社会性的议题或人物塑造也不像是高深而做造势，对爱情故事亦从来无感，看得犯困。

427 不知道为何评分那么高，看了感觉真高辣辣不及basic instinct，高智商犯罪程度不及catch me if you can，对婚姻和爱的探索不及the dogs of babel，逻辑还很混乱。

428 →评论涉及剧透→失望，因果身临其境的双眼有条不紊，罪案反转和婚姻宣告沦为八点档肥皂剧且漏洞频出，残留的日记、监控的访问、如何删除非形象的监控(如删掉，警察没有发现丢失的监控时间?)，婚姻探索太直白，犯罪方式过于庸

429 最佳的

430 王家卫要是没有明星撑场面狗屁都不是，假个球高冷

使用结巴分词对影评语料进行分词，并采用谷歌的开源工具 Word2Vec 将分词的词语表示为词向量。基于 Python 的 Gensim 库中提供了现成的 Word2Vec 工具作为子库。中间结果如下：

```

Pad sequences (samples x time)
[[ [ 0 0 0 ..., 3409 156 21]
[ 0 0 0 ..., 2156 12084 3]
[ 0 0 0 ..., 14 257 3]
...
[ 0 0 0 ..., 81 28 35]
[ 0 0 0 ..., 174 1026 4]
[ 0 0 0 ..., 2 35 3]]

```

将词表示为词向量后，则句子就对应着词向量的集合，也就是矩阵。模型的输入一般只接受一维的特征，因此将矩阵展开，获得一个较长的句向量。

获得句向量后，基于 LSTM 搭建影评情感分析的模型，模型的具体实现是基于 Keras 框架。训练过程如下：

```
X_train shape: (70893L, 80L)
X_test shape: (70892L, 80L)
Build model...
Epoch 1/10
70893/70893 [=====] - 249s - loss: 0.4852 - acc: 0.7653
Epoch 2/10
70893/70893 [=====] - 246s - loss: 0.3686 - acc: 0.8415
Epoch 3/10
70893/70893 [=====] - 246s - loss: 0.3234 - acc: 0.8656
Epoch 4/10
70893/70893 [=====] - 247s - loss: 0.2910 - acc: 0.8816
Epoch 5/10
70893/70893 [=====] - 247s - loss: 0.2662 - acc: 0.8950
Epoch 6/10
70893/70893 [=====] - 247s - loss: 0.2478 - acc: 0.9027
Epoch 7/10
70893/70893 [=====] - 247s - loss: 0.2319 - acc: 0.9098
Epoch 8/10
70893/70893 [=====] - 250s - loss: 0.2185 - acc: 0.9151
Epoch 9/10
70893/70893 [=====] - 251s - loss: 0.2057 - acc: 0.9220
Epoch 10/10
70893/70893 [=====] - 248s - loss: 0.1929 - acc: 0.9272
70892/70892 [=====] - 87s
```

在测试集上的准确率在 85%左右。

通过将模型的输出层设置为 5 个节点，分别对应评论 1 星-5 星。能够更加细粒度地输出影评的情感等级。

3. 实现功能

基于训练好的情感分析模型，对目标功能进行实现。包括：

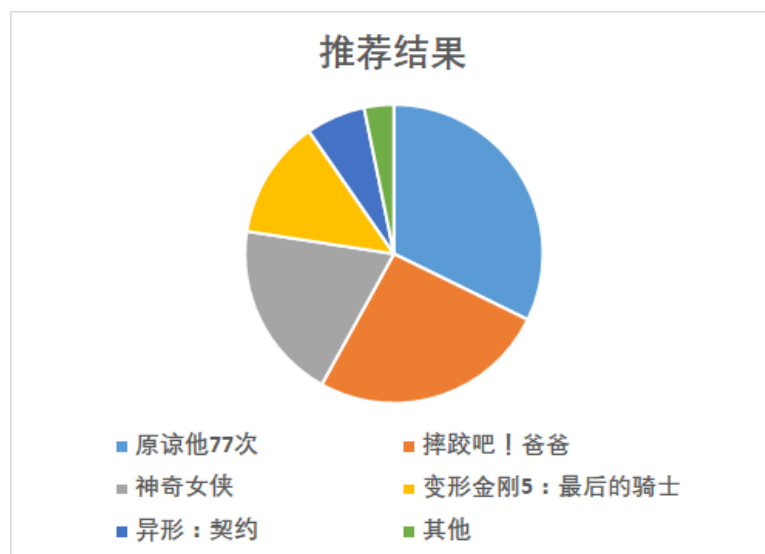
- 用户电影推荐
- 影院电影排片推荐
- 电影种类受欢迎程度评估

3.1. 用户电影推荐

以用户“斯文看客 swain”为例，通过在其主页爬取他看过和评论过的电影，发现其评论过的电影包括《喜欢你》、《华尔街之狼》、《金刚：骷髅岛》等。通过训练后的模型分析其对于不同电影的评论，分类出其对不同电影的评价趋势如下（以部分电影为例）：



可以看出，该用户对于浪漫的爱情电影更为青睐。而我们的模型通过协同过滤算法对于当前院线上映的电影为该用户推荐的结果排名如下：

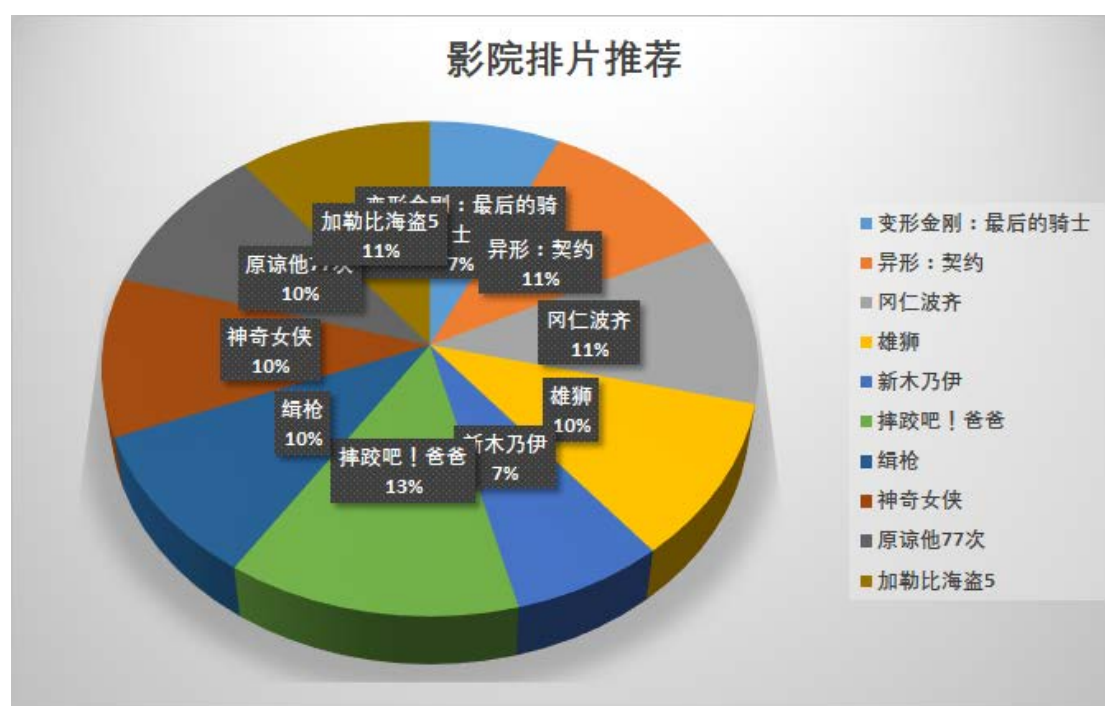


可以看出，推荐的影片第一名为《原谅他 77 次》，其次为《摔跤吧！爸爸》等，可以看出推荐的结果基本符合用户的爱好。

3.2. 影院电影排片推荐

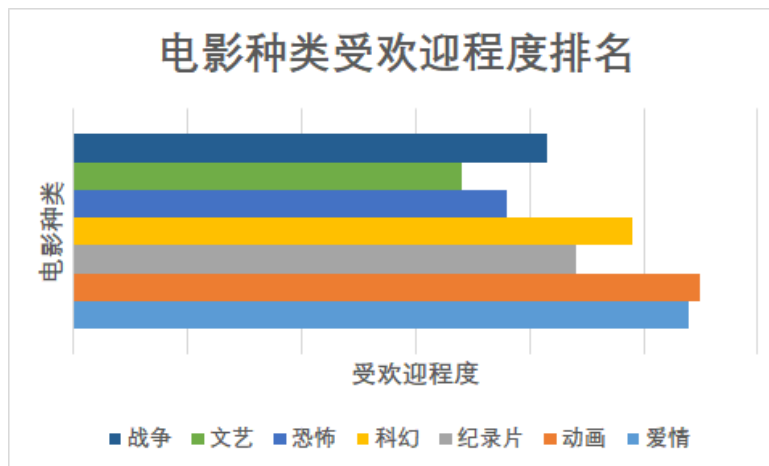
影院的拍片应当以观影者的口碑作为依据，对于口碑好的电影，观影人数较多，影院应当适当地增加排片，从而增加影院的收入。

对于当前院线的电影，通过训练好的模型，对电影的影评进行分析，输出评分的排名，作为电影口碑排名，从而根据这个排名调整影院的拍片。结果如下：



3.3. 电影种类受欢迎程度评估

参考豆瓣电影的电影种类分类，我们从中选取具有标志性的电影种类，包含：爱情类、动画类、纪录片类、科幻类、恐怖类、文艺类和战争类。通过对不同种类电影爬取部分影评，采用训练后的模型进行影评情感分析，计算不同种类电影用户的喜爱程度。并对受欢迎程度进行排名，结果如下：



根据计算结果，电影受欢迎程度排名为：动画>爱情>科幻>纪录片>战争>恐怖>文艺。

总结

随着网络信息的爆炸式增长，从数据中挖掘有价值的内容成为了一个重要的研究问题。社交网络的快速发展，使得人们可以方便地发表自己的想法和观点。对于电影影评，情感代表了用户对于这部电影的直接情感，充分地挖掘这些信息，可以实现很多有意思的功能，例如对用户进行电影的推荐。

本实验实现了一个基于 LSTM 神经网络的影评情感分析系统，并在该系统的基础上实现了几个基于影评文本的数据挖掘任务。神经网络的准确率达到 85% 左右，可以基本完成情感分析任务。同时，基于该系统实现的例如用户电影推荐、影院排片调整推荐等任务，可以从宏观角度地了解用户对于电影的真实情感，从而提升观影者和影院管理者的观影体验以及管理收入。