基于ACSA-RP联合模型对用户评论情感分析与评分预测研究

­­­­

分析得都队

摘 要

随着大语言模型技术与电子商务的快速发展，情感分析在电子商务中越来越受到关注。用户评论背后的情感极性对商业智能有很大的价值。方面类别情感分析(ACSA)和评论评级预测(RP)是检测细到粗的情感极性的两个基本模型，ACSA侧重于预测其在不同方面类别上的潜在情感极性，而RP侧重于从评论内容中预测用户的整体感受，将它们结合起来可以获得更好的分析。为了探究上海市内餐饮店的用户实际体验感受，我们利用python随机爬取了部分上海市内餐饮店的评论数据，包括火锅店、西餐厅、中餐厅等共计49410条评论数据，评论中用户对5个粗粒度（菜品、价格、位置、服务、环境）等多方面进行了直接或含蓄的评价，在5个粗粒度的基础上，我们将评论进一步细化为18个细粒度类别，包括菜品/口味、菜品/外观、价格/性价比、价格/折扣力度、位置/交通方便、服务/上菜速度、环境/就餐空间等类别。在数据预处理的基础上，我们对数据的情感极性打标签，最终的语料库共由46864条真实评论组成，我们将其随机分为训练集（36922）,验证集（4971）和测试集（4971）。

我们通过建立ACSA-RP联合学习模型，我们进行了一系列实验来评估ACSARP联合模型的效果。我们采用了BERT-Base Chinese预训练模型，冻结了BERT模块的参数，并调整了训练参数以优化模型性能。实验结果显示，在训练过程中，情感分类准确率逐步提升，评分回归RMSE逐渐下降，最终稳定在70%左右的分类准确率和0.7左右的评分RMSE。在整个数据集上评估，训练集上的准确率为76%，RMSE为0.7009，测试集上的准确率为73%，RMSE为0.7039。进一步分析显示，模型对积极类和中性类预测效果较好，但对消极类预测效果一般，可能是由于消极类样本数过少。注意力汇聚示例展示了模型在关注与目标方面类别最相关的词元方面的表现，进一步证明了联合模型的有效性。

本研究的情感分析结果为餐饮业管理者提供了深入的消费者心理和偏好洞察。通过分析顾客评论，餐厅可以调整菜单、优化服务、改善顾客体验。具体建议包括菜品口味优化、服务质量提升、环境氛围改善、市场营销、预测分析等。我们利用ACSA-RP联合模型对上海市餐饮服务进行了深入的情感分析，从而理解了顾客满意度和关键影响因素。这些发现不仅可为餐饮业提供改进策略，也为其他服务行业提供参考。随着技术的进步，未来可进一步细化情感分析模型，以创新方式满足消费者需求，推动服务行业发展。

一、背景介绍与研究问题

1.1背景介绍

随着数字化经济的加速发展，某团、某点评等在线评价平台已成为公众选择餐饮服务的重要参考。这些平台每日积累大量的用户生成内容（UGC），如评分、文字评论等，这些数据不仅反映了消费者的直接体验，也构成了餐厅服务质量的即时反馈。在竞争日益激烈的餐饮市场，如何有效利用这些在线评论来改善服务和提高顾客满意度，是餐饮业者亟待解决的问题。

然而，评论的高度主观性和复杂性使得从中提取有用信息成为一项挑战。传统的情感分析技术在处理如此复杂和多样的数据时常常力不从心，无法准确捕捉到评论中的细微情感变化和具体方面的情感倾向。因此，发展高级的自然语言处理（NLP）技术，尤其是深度学习和预训练模型（如BERT），对于精确解析用户评论具有重要意义。这些技术可以在更细的粒度上，如方面级（Aspect-Based）情感分析，对评论内容进行深入解析，识别出消费者对餐厅食物口味、服务态度、环境氛围等具体方面的情感反应。

本研究利用ACSA-RP联合模型的NLP技术，通过深层次的文本分析，不仅可以揭示消费者的总体情绪倾向，还能详细地解读消费者对餐厅各个服务方面的具体评价，从而为餐厅管理者提供了一种全新的、基于数据的服务改进方法。这种方法的实施，有助于餐厅精确地调整其商业策略，针对性地改善服务质量，最终达到提升顾客满意度和增强顾客忠诚度的目的。

1.2研究问题

本研究旨在通过深入分析某点评上的用户评论，探索消费者情感的具体表现及其对餐厅业务的实际影响。鉴于在线评论的丰富性和复杂性，本研究设计了几个关键的研究问题来系统地解析和理解这些数据。首先，我们将探讨评论中情感倾向的总体分布情况，包括积极、中性和消极情感的比例，以及这些情感如何分布在不同类型的餐厅和不同的服务方面。研究将深入分析不同餐厅类型（如快餐店、家庭式餐厅、高端餐厅等）在顾客评价中的情感特点。这一分析帮助我们了解在不同市场定位的餐厅中，消费者的期望和满意度有何不同，并探讨哪些服务方面的改进最能提升顾客的整体满意度。

此外，我们还将通过方面级情感分析来详细探讨顾客对餐厅各个具体方面（如食品口味、服务质量、环境舒适度等）的情绪表达，识别出对顾客满意度影响最大的因素。这些细致的分析将为餐厅提供具体的改进建议，使其能够在正确的方向上努力，更有效地满足顾客需求。最后，本研究期望通过对这些复杂数据的系统分析，建立情感分析结果与餐厅实际经营状况之间的联系，为餐饮业者提供改进服务的具体方向，同时也希望能为其他服务行业提供利用在线评论进行质量控制和服务改进的参考。

二、数据的说明与描述

本报告目的是对某点评上的评论背后蕴含的情感含义进行分析，以此来对餐饮店的真实经营情况进行刻画。本报告模型建立的数据主要来源于某点评平台网页，利用python随机爬取了部分上海市内餐饮店的评论数据，包括火锅店、西餐厅、中餐厅等共计49410条评论数据，在评论中用户对5个粗粒度（菜品、价格、位置、服务、环境）等多方面进行了直接或含蓄的评价。在此基础上，我们首先对数据进行预处理，包括（1）删除用户隐私信息（例如用户id、用户名、头像和上传时间）；（2）过滤掉50字以内的短评论和1000字以上长评论；（3）过滤掉非汉字比例超过70%的评论；（4）过滤掉重复字符连续超过5个以上的低质量评论。

用户发布的评论中已包含总体五星评分和评价内容，我们在数据预处理的基础上对数据的情感极性打标签，最终语料库共由46864条真实评论组成，我们将其随机分为训练集（36922）,验证集（4971）和测试集（4971），并将5个粗粒度类别细化为18个细粒度类别，具体请见表1。

表1 18个方面类别和定义的完整列表

|  |  |  |  |
| --- | --- | --- | --- |
| 方面类别 | 定义 | 方面类别 | 定义 |
| 菜品/口味 | 菜品味道 | 位置/容易寻找 | 餐厅是否容易寻找 |
| 菜品/外观 | 菜品品相 | 服务/排队时间 | 排队时间是否可接受 |
| 菜品/分量 | 菜品分量 | 服务/顾客态度 | 餐厅服务态度 |
| 菜品/推荐程度 | 菜品是否值得被推荐 | 服务/停车方便 | 餐厅是否容易停车 |
| 价格/消费水平 | 消费等级 | 服务/上菜速度 | 上菜速度快慢 |
| 价格/性价比 | 菜品是否性价比高 | 环境/装修情况 | 餐厅装修等级 |
| 价格/折扣力度 | 折扣大小 | 环境/嘈杂情况 | 餐厅噪音等级 |
| 位置/中心商圈 | 餐厅是否位于中心商圈 | 环境/就餐空间 | 用餐空间和座位大小 |
| 位置/交通方便 | 餐厅是否交通方便 | 环境/卫生情况 | 餐厅卫生情况 |

三、数据的分析与解读

我们首先根据上一节细化为18个细粒度方面类别的定义对语料库清洗后的数据绘制总体分布图，通过积极、中性和消极评论数量对应的柱体高度展示各类别的情感总体偏向程度，具体见图1。

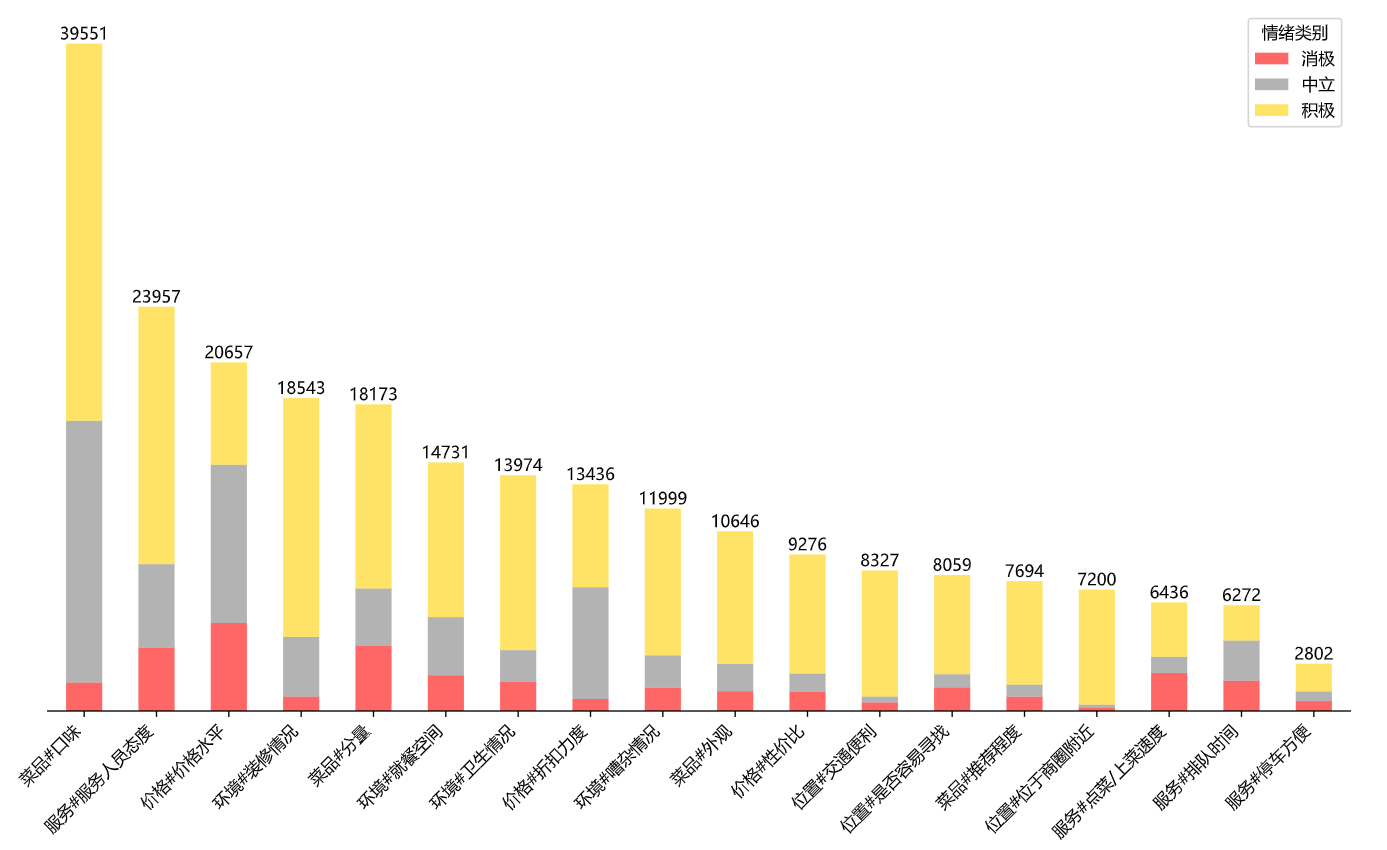


图1 语料库中18个细粒度方面类别的分布情况与情感总体偏向程度

为了探索星级评分与评论语句之间的复杂关系，我们首先探索了不同评论字数下的不同星级评分情况，具体见图2。

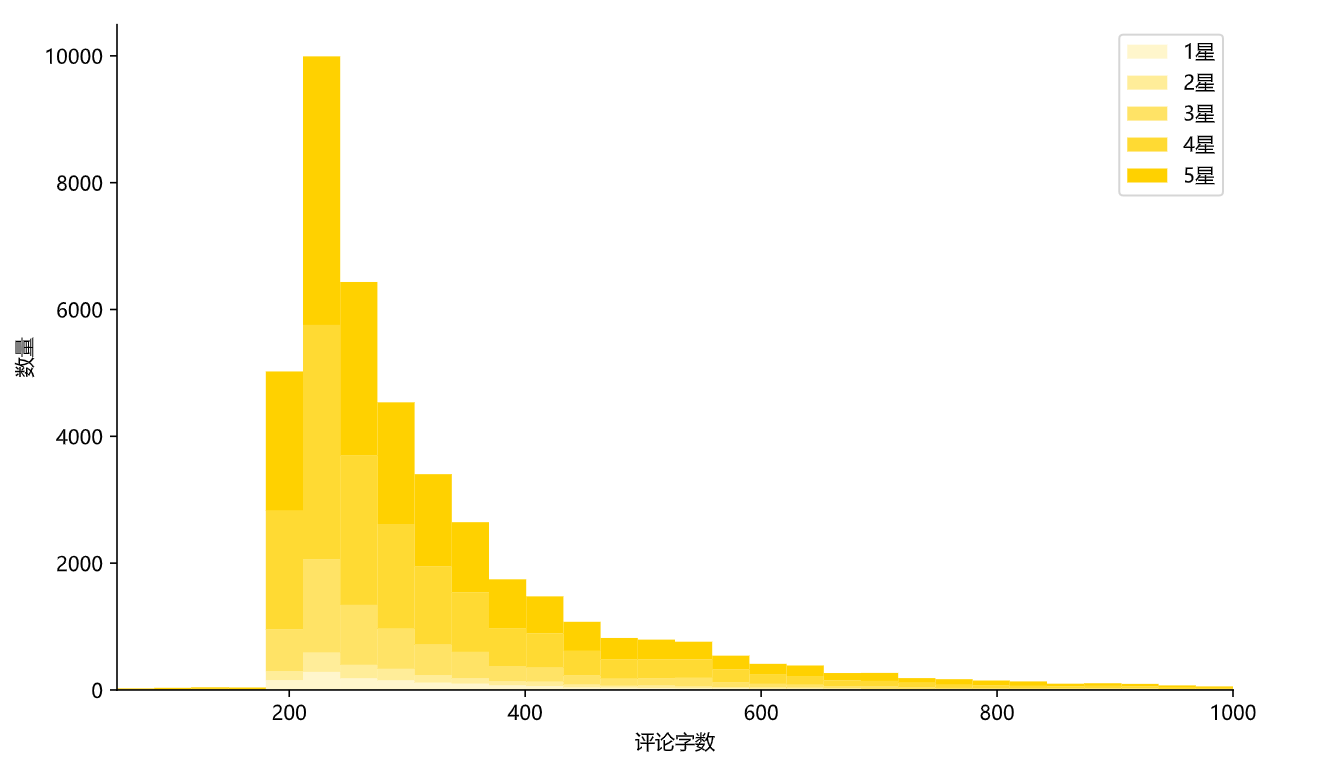


图2 语料库中评论字数与星级评分总体分布情况

通过图表可以直观分析出评论字数在整体上集中在200-400字之间，星级分布情况集中在3星-5星之间。进一步地，当给定一个评论时，我们对其的18个细粒度方面类别情感极性分别标记为1(积极)，0(中立)或-1(消极)；但数据会出现表达混淆模糊无法判断情感的情况，我们需要对疑难案例进行注释指南，以下为几种典型案例，具体见表2。

表2 几种典型疑难案例注释指南

|  |  |  |  |
| --- | --- | --- | --- |
| 类别 | 例子 | 指南 | 情感极性 |
| 情感时移 | 我之前挺喜欢这家餐厅的饭菜，不过今天的饭菜可不怎么样。 | 当评论中存在随时间移动的情感时，采用最近的情感极性。 | （菜品/口味，-1） |
| 情感含蓄 | 比五星级酒店的餐厅差远了，而且在五星级酒店中餐厅里两个人吃一顿也就500左右就够了。 | 当评论情感过于含蓄，需要判断对比等修辞情况，采用隐式情感极性标注。 | （价格/消费水平，-1） |
| 意见冲突 | 这道菜有点咸，但是味道很赞。 | 当对同一方面类别存在多个情感极性时，选择优势情感。 | （菜品/口味，1） |
| 情感温和 | 饭菜还可以，不过也算不上特别好吃。 | “中立”的标签适用于温和偏积极或偏消极的情感。 | （菜品/口味，0） |
| 无关干扰 | 上次去的一家店很难吃，今天来了这家新的，感觉很好吃。 | 评论提到了顾客过去打卡的餐厅，但我们只关注被评价的餐厅 | （菜品/口味，1） |

在本报告中，我们考虑的模型是ACSA与RP，它们在电子商务中的商业智能应用都是很广泛的，并且是高度相关和互补的。ACSA侧重于预测其在不同方面类别上的潜在情感极性，而RP侧重于从评论内容中预测用户的整体感受，将它们结合起来可以获得更好的分析。数据在18个细粒度方面类别被标记情感极性之后，数据可以较好地符合ACSA模型与RP模型的预测效果，提高预测精度。

四、模型的建立与求解

4.1 ACSA-RP联合模型的建立

我们用表示训练数据中用户评论语料库的集合。给定一个由一系列单词组成的评论，首先ACSA模型旨在预测评论对于上述18个方面类别的3个不同的情感极性。表示评论的长度，为预定义的方面类别的数量(即本报告中为18个)。假设评论中有个提到的方面类别。我们定义向量来表示方面类别的出现情况。当在评论中提到方面类别时，，否则,所以有。对于RP模型，它可以预测的5星级评分，其中代表给定评论的总体评分。

对于一系列给定的用户评论，ACSA侧重于预测其在不同方面类别上的潜在情感极性，而RP侧重于从评论内容中预测用户的整体感受，将它们结合起来可以获得更好的分析。评论评级预测可以看作是一个单句分类回归任务，基于于此我们提出了ACSA-RP联合学习模型，以多任务学习的方式来解决ACSA和RP，我们的联合模型采用了BERT编码器从细到粗的语义表示能力。图3展示了联合模型的框架。

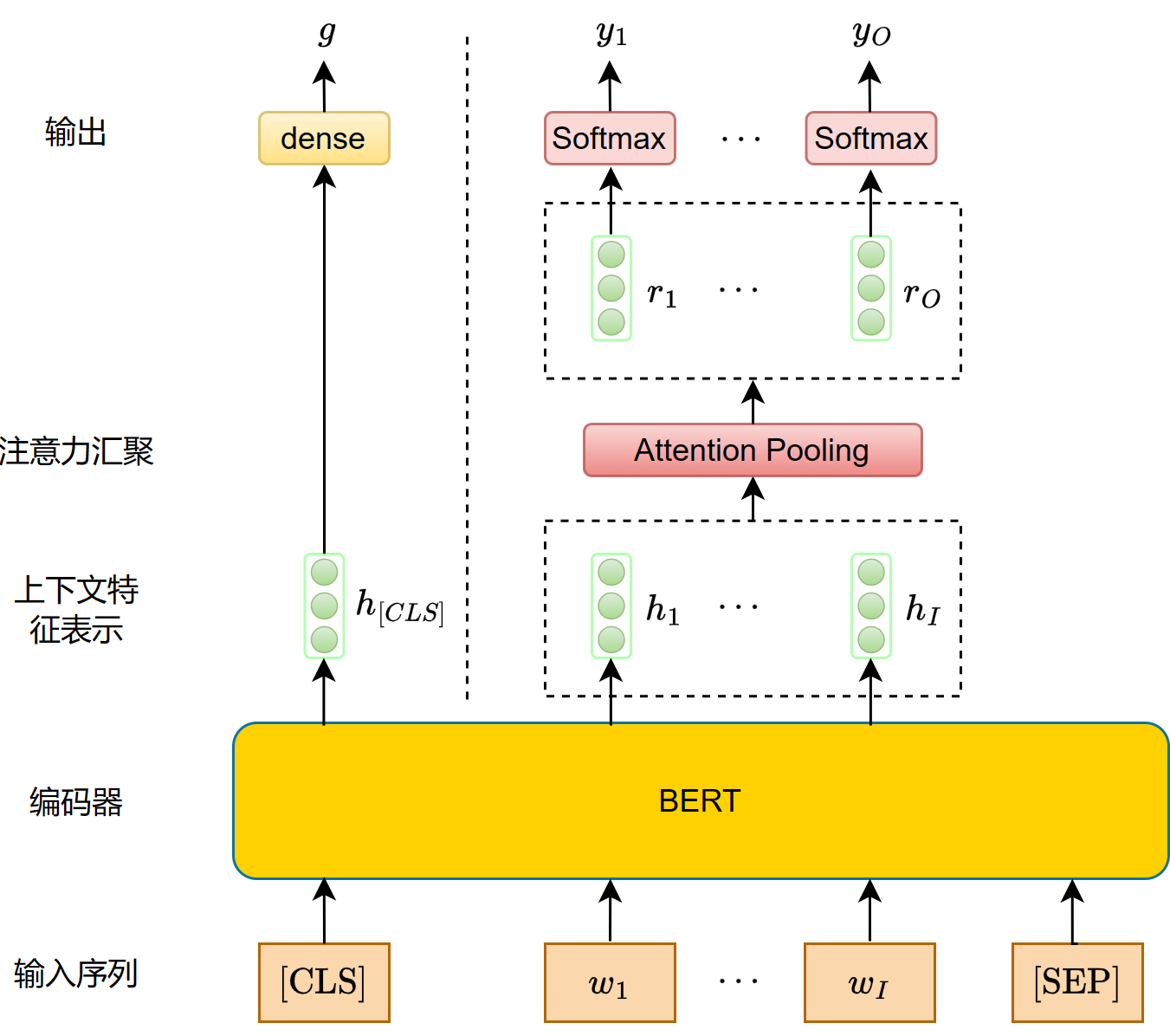


图3 ACSA-RP联合学习模型框架

在ACSA-RP联合学习模型框架中，虚线的右侧部分用于预测多个方面类别的情绪极性，左侧部分用于预测评论评分。输入评价的嵌入标签是通过共享BERT编码器生成的，设是特征嵌入向量，其中是隐藏层的大小，是给定评论的长度。由于不同的方面类别信息分散在评论中，我们添加了一个注意力汇聚层来动态聚合每个方面类别的相关特征嵌入。注意力汇聚层将会在模型中筛选与目标方面类别最相关的特征。

其中。是一个由所有标记的注意力权值组成的向量，这些标记会选择与方面类别相关的标记区域，是关于第方面类别的评论注意力表征。然后我们有

其中和是softmax层的训练参数，是标签的数量（在本报告中是3个情感极性）。

ACSA-RP联合学习模型的ACSA损失函数定义为：

如果S中没有提到方面类别，则将设为随机值。作为门函数，过滤掉随机的，保证只有评论中提及的方面类别参与损失函数的计算。

对于评分预测部分，我们采用BERT编码器生成的作为输入评论的嵌入表征，其中为BERT编码器中隐藏层的大小,

在给定评论的情况下，ACSA-RP联合学习模型中RP损失函数定义为：

其中是可训练的参数。

因此ACSA-RP联合学习模型的最终损失函数为：

4.2 ACSARP联合模型的求解

我们开展了一系列实验来评估ACSARP联合模型的效果。

**实现细节：**BERT模型采用预训练模型 [BERT-Base Chinese](https://huggingface.co/google-bert/bert-base-chinese)，训练时冻结BERT模块的参数，学习评分预测任务中的全连接层参数以及情感分类任务中的注意力汇聚层参数，总参数量为21M；Batch size 设置为64；优化器选用Adam(Kingma and Ba, 2014)，相应参数 ；最大序列长度设置为512；总训练epoch数设置为20；初始学习率设置为0.00005；实验环境所用显卡为NVIDIA GeForce RTX 3060。另外，考虑到每个方面情感分类中存在的样本类别不均衡的情况（如：多数评论中不含有对“服务/停车方便”的评价，以及已有的评价里“积极”情感类多于其他两类），损失函数部分采用了带权的[交叉熵损失](https://pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html)，权重为每个类别三类样本频数的倒数，而在计算准确率时，我们考虑的是非缺失值（标签不为-2）的预测准确率

**结果分析：**下图展现了训练过程中情感分类准确率和评分回归RMSE的变化。可见随着迭代进行，在单个batch上的分类准确率逐渐提升，评分RMSE逐渐下降。最终分类准确率稳定在70%左右，评分RMSE稳定在0.7左右。我们也在整个数据集上评估了模型表现，得到在训练集上的准确率为76%，RMSE为0.7009，在测试集上的准确率为73%，RMSE为0.7039。

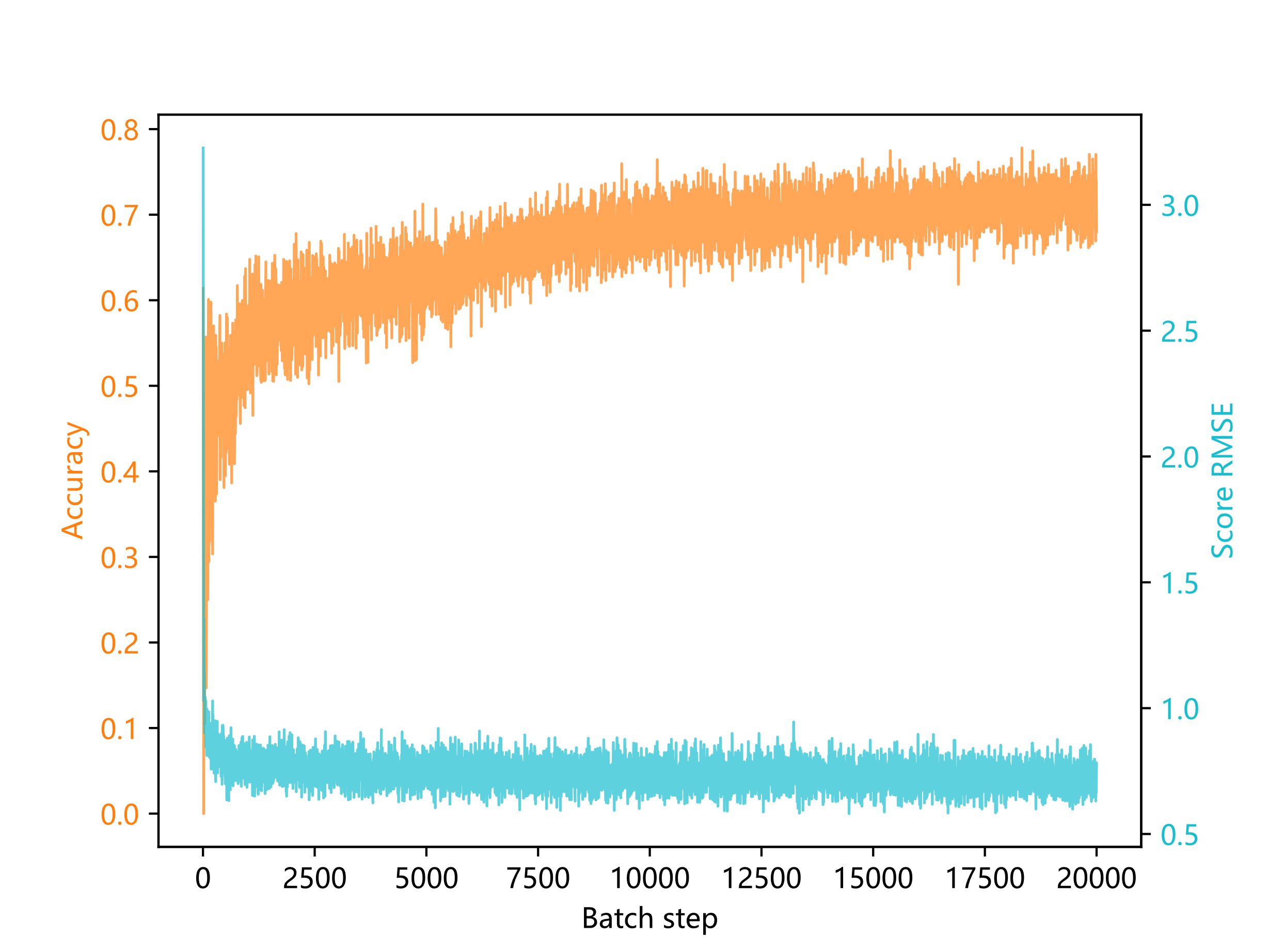


图4 训练过程中准确率和评分RMSE变化

为了进一步分析模型结果，对于评分预测任务，我们绘制了预测评分与实际评分之差的直方图，对于预测任务，绘制了“菜品/口味”类的混淆矩阵。在直方图中，观察到误差总体聚集在0附近，但中心略向-1偏移，表明模型预测评分低于实际评分。在混淆矩阵热力图中（横轴表示实际标签，纵轴表示模型预测分类），观察到模型对积极类和中性类预测效果较好，对消极类预测效果一般，可能原因是样本不均衡，消极类样本数过少。

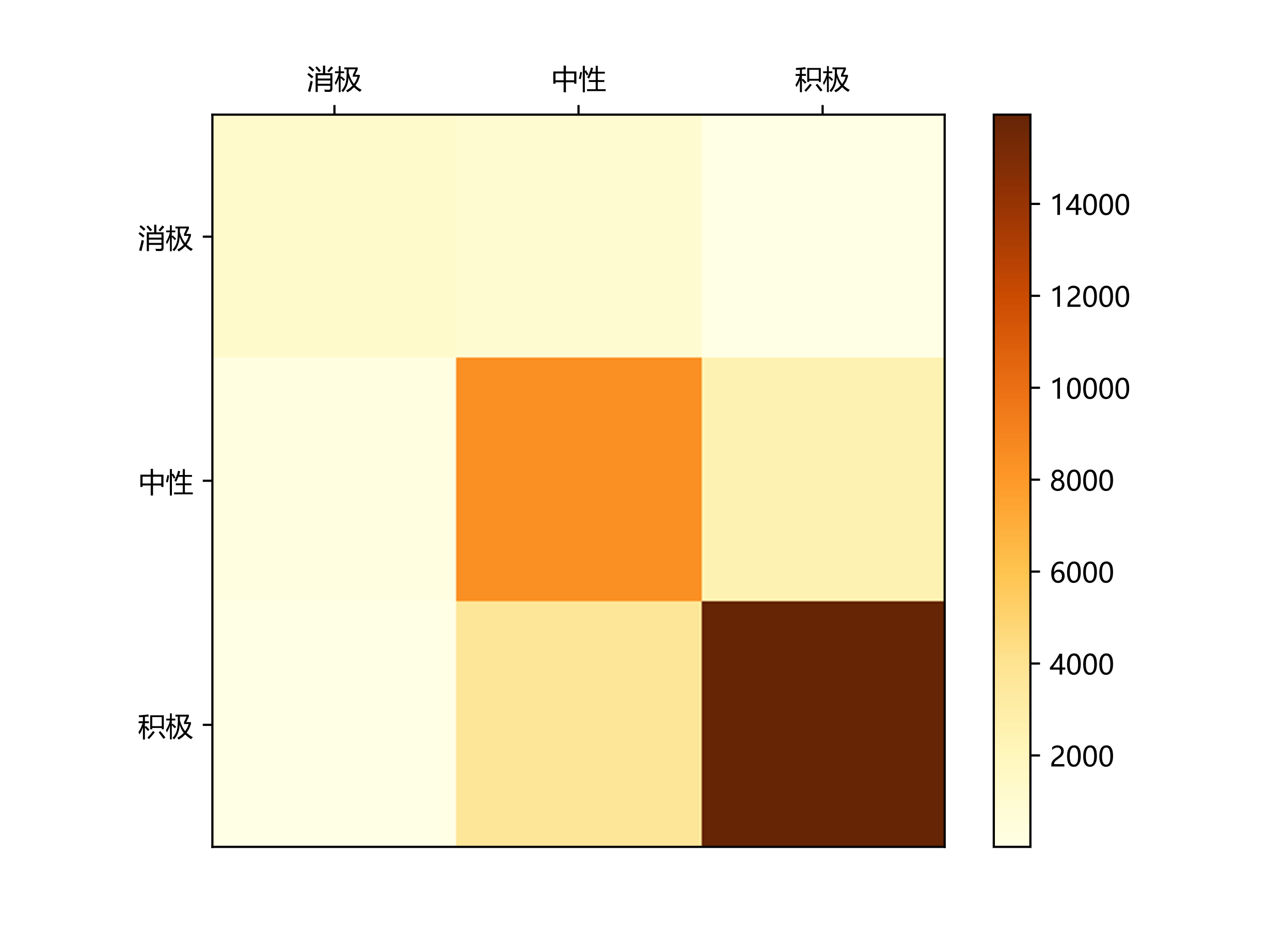
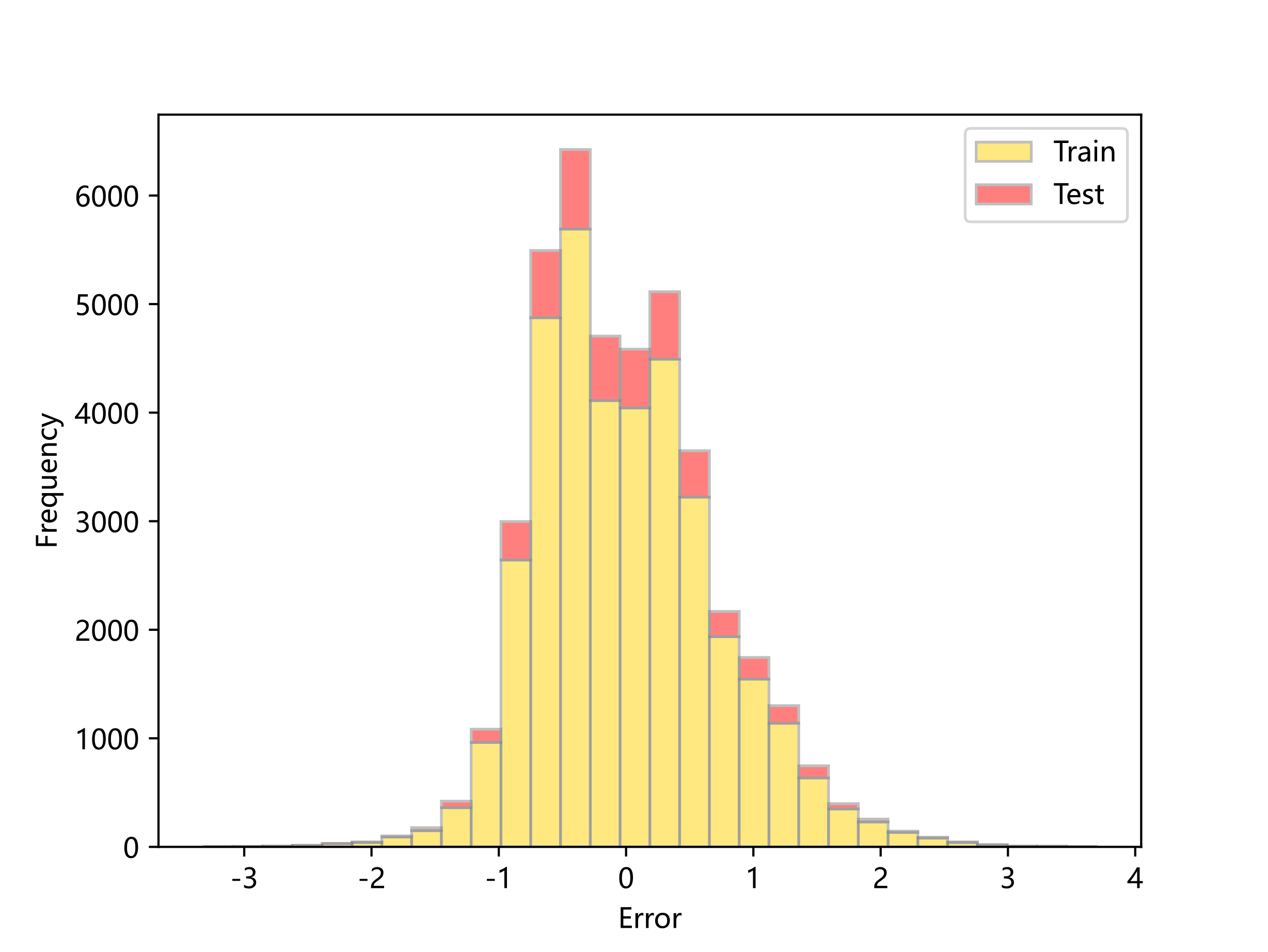


图5 RMSE直方图和混淆矩阵热力图

**注意力汇聚示例：**为了更好地展示注意力汇聚在模型中的作用，我们将模型的注意力权重可视化。 由于不同方面评价的信息分散在评论中，我们添加了一个注意力汇聚层来动态地为每个方面类别聚合相关的词元嵌入。注意力汇聚层帮助模型关注与目标方面类别最相关的词元。下图展示了三个给定方面类别的注意力权重。颜色的强度代表注意力权重的大小，这意味着词元与给定方面类别的相关性。很明显，我们的联合模型能够关注到评论中与方面类别最相关的词元。

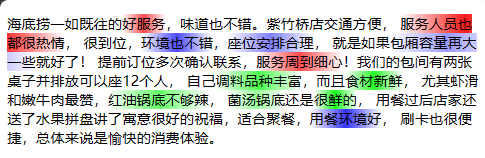


图6 注意力汇聚示例

五、商业应用与总结

5.1商业应用

本研究的情感分析结果对于餐饮业管理者提供了深入的消费者心理和偏好洞察。通过系统地分析顾客评论中的情感倾向，餐厅可以更有效地调整菜单，优化服务流程，以及改善顾客体验。以下是几个具体的商业应用建议：

1.菜品和口味优化：情感分析显示，顾客对某些菜品的满意度较高，而对其他菜品的评价不佳。餐厅可以依据这些信息调整菜单，重点推广受欢迎的菜品，并改进或替换评价较低的菜品。

2.服务质量提升：通过分析关于服务的评论，餐厅可以识别服务中的问题点，如服务速度慢、态度不佳等。据此，餐厅管理者可以实施针对性的员工培训程序，提高服务人员的服务意识和效率。

3.环境与氛围改善：消费者对餐厅环境的评价直接影响其就餐体验。本研究结果可帮助餐厅了解顾客对环境的具体需求，从而进行相应的环境布置和氛围营造。

4.市场营销和促销活动：通过分析顾客对价格和促销活动的反馈，餐厅可以设计更吸引人的营销策略和促销活动，以吸引新顾客并保持老顾客的忠诚度。

5.预测分析：结合情感分析的结果和其他业务数据（如销售数据），餐厅可以建立预测模型，预测顾客流量和销售趋势，以更好地管理库存和人力资源。

5.2总结

本研究利用ACSA-RP联合模型的自然语言处理技术对某点评上的用户评论进行了深入的情感分析，以揭示消费者对上海市餐饮服务的详细看法。通过分析评论中的情感倾向和内容细节，我们不仅理解了顾客对各方面服务的具体满意度，还识别出了影响顾客体验的关键因素。这些发现对于餐饮业者来说极具价值，可以用来优化服务、调整菜单和改善顾客互动。

通过对18个细粒度方面类别的情感分析，我们得以详细探索和量化了顾客关于食品口味、服务质量、环境氛围等方面的情感反应。。这样的具体反馈使得餐厅能够针对性地进行改进，比如通过增加人手或优化工作流程来缩短等待时间。本研究的方法论为餐厅管理者提供了一种系统的方式来理解和利用在线评论。通过实施基于数据的决策制定，餐厅不仅可以提高顾客满意度，还可以增强其市场竞争力。

本研究的成果不仅为餐饮业提供了改进服务的具体策略，也为其他服务行业如零售、酒店等提供了如何利用消费者评论来优化服务的参考。随着技术的进步，未来的研究可以进一步细化情感分析模型，例如通过集成更多的语境和语言特征来提高情感识别的准确性。通过这项研究，我们强调了在快速变化的消费市场中，理解和应用消费者反馈的重要性。我们希望这种基于数据的洞察能够激励更多的企业采纳科技驱动的方法，以创新的方式来满足消费者需求，并推动服务行业的持续发展。

参考文献

[1] Zehui Dai, Wei Dai, Zhenhua Liu, Fengyun Rao, Hua-jie Chen, Guangpeng Zhang, Yadong

Ding, and Jiyang Liu. 2019. Multi-task multi-head attention memory network for fine

grained sentiment analysis. In CCF International Conference on Natural Language Processing and Chinese Computing, pages 609–620. Springer.

[2] Gayatree Ganu, Noemie Elhadad, and Am´elie Marian. 2009. Beyond the stars: improving rating predictions using review text content. In WebDB, volume 9, pages 1–6. Citeseer.

[3] Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In Proceedings

of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining,

pages 168–177.

[4] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. arXiv

preprint arXiv:1412.6980.

[5] Fangtao Li, Nathan Nan Liu, Hongwei Jin, Kai Zhao, Qiang Yang, and Xiaoyan Zhu. 2011.

Incorporating reviewer and product information for review rating prediction. In Twenty-

second international joint conference on artificial intelligence.

[6] Dehong Ma, Sujian Li, Xiaodong Zhang, and Houfeng Wang. 2017. Interactive attention

networks for aspect-level sentiment classification. In International Joint Conference

on Artificial Intelligence, pages 4068–4074.

[7] Maria Pontiki, Dimitrios Galanis, John Pavlopoulos, Haris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. Semeval-2014 task 4: Aspect based sentiment analysis. In International Workshop on Semantic Evaluation (SemEval 2015).