

附件 3-1

项目编号	
------	--

天津市大学生创新训练计划项目申报书

学 校 名 称 : 天津工业大学

基 于

RIME-CNN-LSTM-Attention

项 目 名 称 : 的平台用户购买行为研究

项 目 级 别 : ☒ 国家级 ☐ 市级 ☐ 校级

所属专业类代码: 0701-数学

项 目 负 责 人 : 李晓达

联 系 电 话 : 18860582417

指 导 教 师 : 刘明

联 系 电 话 : 13821627465

申 报 日 期 : 2024 年 4 月 21 日

天津市教育委员会
二〇一九年一月

填写说明

一、项目申报书要逐项认真填写,填写内容必须实事求是,表达明确严谨。空缺项要填“无”。

二、每个项目参与学生 3-6 人,指导教师不得超过 2 人。

三、“项目编号”由学校统一填写。编号规则:2019+高等学校代码(如:天津科技大学—10057)+3 位流水号。

四、项目所属专业类代码:四位代码,按照《普通高等学校本科专业目录和专业介绍(2020 年)》填写。

五、项目级别为国家级、市级、校级。

六、项目期限一般为 1-2 年。

七、格式要求:表格中的字体采用五号宋体,单倍行距;需签字部分由相关人员以黑色钢笔或签字笔签名。

项目名称		基于 RIME-CNN-LSTM-Attention 的平台用户购买行为研究					
项目所属专业类代码		0701-数学		项目级别		国家级	
财政拨款（元）				学校拨款（元）			
项目总经费（元）		20000					
项目实施时间		起始时间： 2024 年 5 月 完成时间： 2026 年 5 月					
申请人或申请团队		姓名	年级	学号	所在院系/专业	联系电话	E-mail
	负责人	李晓达	2023 级	2310810085	数学类	18860582417	3150473133@qq.com
	成员	徐悦玲	2023 级	2310810010	数学类	18500677130	3045529282@qq.com
		孟祥成	2023 级	2310810037	数学类	15030980950	3130480371@qq.com
		杨文婷	2023 级	2310810160	数学类	15222611078	3302792328@qq.com
		王薪颖	2023 级	2310810168	数学类	15808043526	15808043526@qq.com
		徐嘉怡	2023 级	2310810053	数学类	18822162533	1368461006@qq.com
指导教师	第一指导教师	姓名	刘明		专业技术职务	副教授	
		单位	数学科学学院		研究方向	智能算法与科学计算	
		联系电话	13821627465		E-mail	liuming@tiangong.edu.cn	
	第二指导教师	姓名			专业技术职务		
		单位			研究方向		
		联系电话			E-mail		

一、项目简介（200 字以内）

随着互联网技术的发展，社交媒体已经成为人们日常生活中不可或缺的一部分。用户在社交媒体上的互动、分享等行为产生了大量的数据，这些数据蕴含着用户的兴趣、需求等信息。因此，分析社交媒体用户的购买行为，对于企业制定营销策略、优化产品和服务具有重要意义。用户购买行为受到多种因素的影响，它们随时间变化并相互影响，形成了复杂的多变量时间序列数据。如何从这些数据中提取有效信息并进行准确预测，是当前数据分析领域面临的一个挑战。

二、申请理由（包括自身具备的知识条件、自己的特长、兴趣、已有的实践创新成果等）

- 1. 专业优势：**项目组全体成员均为数学专业的学生，具有扎实的数理知识基础，并系统学习了编程语言、数据结构等相关课程。
- 2. 技能优势：**项目组成员中三名成员熟练掌握 MATLAB, Dev-C++ 编程软件操作，能够使用 SPSS, SQL 等软件处理数据，具有良好的编程能力和数据分析能力。两名成员参加过亚太杯、美赛等数学建模竞赛。三名成员熟练掌握 Python 并熟悉 Python 深度学习内容，全员学习并掌握 C 语言。两名成员通过计算机《二级 MS Office 高级应用》熟练掌握 word, excel, ppt, ps, pr 等技术软件。全体成员均学习成绩优异，均在本科第一学期高分通过了英语四级考试，多名成员参加全国大学生英语竞赛，国青杯翻译大赛，韩素音翻译大赛，词达人比赛。
- 3. 特长优势：**成员们具有扎实的阅读英文文献和检索的能力，能够获取研究领域的前沿信息与成果。成员对于学习计算机编程报以极大的热忱，对学习新知识有很高的主动性和积极性，经常询问老师、与同学讨论。

三、项目方案

具体内容包括：

1.项目研究背景（国内外的研究现状及研究意义、项目已有的基础，与本项目有关的研究积累和已取得的成绩，已具备的条件，尚缺少的条件及方法等）；

1.1 国内外研究现状及研究意义

目前对用户购买行为的预测主要包括点击率、购买行为、用户活跃度等预测，用户复购行为预测也是购买行为预测的一种。传统的购买行为预测模型包括 S M C 模型、R F M 模型和 B G / N B D 模型等，多数学术研究者都是在这三大模型的基础上作出改进，以得到更好的预测效果。目前随着人工智能的兴起，学术研究者多数将大数据分析技术与机器学习相结合，不断有学者尝试使用深度学习、神经网络、随机森林、XGBoost 等模型解决此类问题，以求得到更高的准确度和精准性，并使用集成技术将多个分类器融合在一起来提高模型性能。基于 RIME-CNN-LSTM-Attention 的多变量回归预测模型结合了卷积神经网络（CNN）、长短期记忆网络（LSTM）和注意力机制，能够捕捉数据中的复杂模式，深入了解用户行为并能够精准预测购买意向，进而优化营销策略。此外，动态特征捕捉能力能够捕捉用户行为的动态变化，实时分析用户需求，提升用户体验，并为企业战略提供依据，最终提升竞争优势。

1.2 项目已有的基础

查阅大量国内外相关研究文献，了解研究项目如今的发展现状，以及前人所运用的研究方法。询问财经专家和数学专家。学习并掌握研究项目方面相关的知识，了解当今社会企业所欠缺的以及迫切需要提高或优化的方面。确定研究目标，即预测平台用户的购买行为、发现潜在客户来提高销售转化率和洞察市场趋势。制定具体的研究方法和过程，为后期项目的开展与研究打下基础。

1.3 项目积累

近年来，随着网购的普及，计算机技术和人工智能技术的快速发展，算法模型在网络购物等方面的地位也越来越重要。RIME（霜冰优化算法）、CNN（卷积神经网络）、LSTM（长短期记忆递归神经网络）和 Attention 是目前研究用户购买的过程中最具发展前景的模型算法：RIME-CNN^[1]是一项 2023 年发表于 SCI、中科院二区 Top 期刊《Neurocomputing》上的新型优化算法；积神经网络具有良好的特征提取能力和分类能力，可以自动学习，在研究用户购买行为上拥有较为重要的地位；LSTM^[2]可用于解决一般递归神经网络中普遍存在的长期依赖问题，使用 LSTM 可以有效的传递和表达长时间序列中的信息并且不会导致长时间前的有用信息被忽略；Attention^[3]可以一步到位获取全局与局部的联系，且参数少，模型复杂度低，在算法优化方面有较为重要的意义。

[1]李思涌,吴书汉,孙伟.基于注意力机制的 CNN-LSTM 网络车内 CAN 总线入侵检测技术[J].信息安全研究,2023,9(10):961-967.

[2]李家蕊,杨旻.基于 LSTM-LightGBM 模型的烟草存销比层次化预测方法[J/OL].烟台大学学报(自然科学与工程版):1-7.DOI:10.13951/j.cnki.37-1213/n.230808.

[3]周祥,张世明,苏林鹏,张守平.基于注意力机制与 LSTM-CCN 的月降水量预测[J/OL].人民长江:1-9.

2. 项目研究目标及主要内容

2.1 研究目标

1) 理解用户需求，预测其购买行为：先深入了解用户在社交媒体上的需求和偏好，为个性化推荐提供依据，然后通过模型准确预测社交媒体用户的购买意向和实际购买行为。

2) 发现潜在客户，提高销售转化率：识别具有购买潜力的用户，为企业拓展市场提供支持，从而助力企业将更多的潜在用户转化为实际购买者。

3) 洞察市场趋势，提升竞争优势：从大量用户数据中洞察市场趋势和变化，为企业决策提供参考，并利用先进的模型技术，提升企业在市场中的竞争优势。

2.2 主要内容

1) 收集社交媒体数据并提取特征：获取用户在社交媒体上的相关数据，并从收集的数据中提取与购买行为相关的特征。

2) 构建、训练、优化并评估模型：利用 RIME-CNN-LSTM-Attention 构建多变量回归预测模型，通过训练数据对模型进行训练，并优化模型的性能，最后采用合适的指标评估模型的预测准确性。

3) 对模型结果进行相关分析并得出结论：首先深入分析用户在社交媒体上的行为模式，接着探究影响用户购买决策的关键因素，其次考虑用户行为的时间序列特征，然后研究用户在社交网络中的关系对购买行为的影响，最后比较不同社交媒体平台上用户购买行为的差异；最终得出结论并根据用户特征进行分群，为个性化营销提供依据。

3. 项目创新特色概述

3.1 特色

理论创新 ➡ 结合了雾凇算法优化的卷积神经网络（RIME-CNN）和基于注意力机制的长短记忆网络（LSTM-Attention）

成果创新 ➡ 可以有效地处理和分析社交媒体用户购买行为中的复杂时间序列数据，从而提高预测的准确性和可靠性。

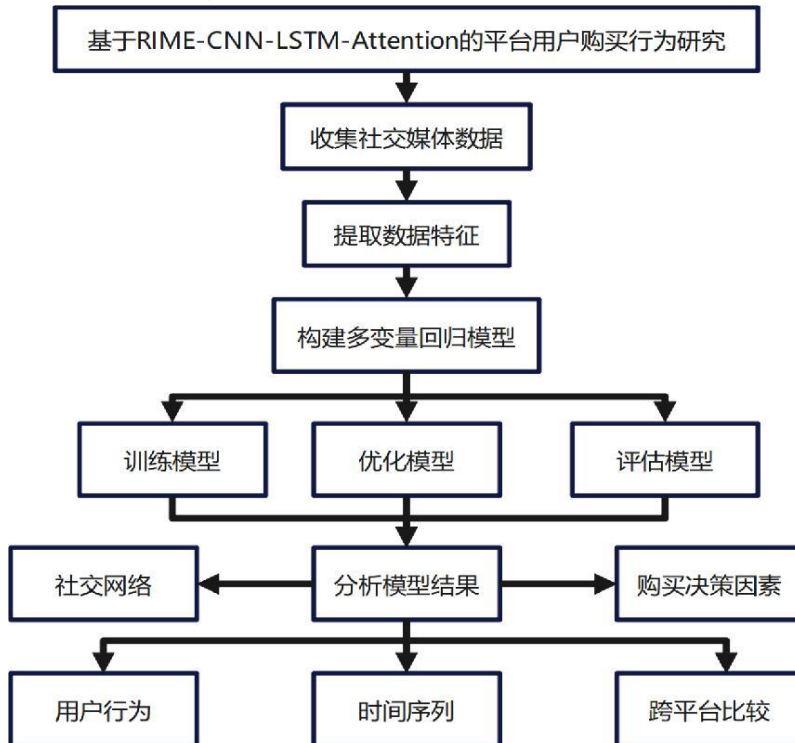
3.2 创新

1) 多模态融合：结合了多种不同的模型，雾凇算法（RIME）、卷积神经网络（CNN）、长短期记忆网络（LSTM）和注意力机制，能够充分利用数据的不同特征。

2) 泛化能力：在不同的社交媒体平台和用户群体中具有较好的泛化能力。

3) 多变量处理能力：可以处理多个相关变量，更好捕捉用户购买行为的复杂关系。

4. 项目研究技术路线



5. 研究进度安排:

第一步，课题准备阶段

(2024 年 4-10 月)

- (1) 依据初步研究框架，收集整理相关文献资料，了解国内外相关研究进展。
- (2) 参与有关交流会议以及讲座，对知识以及知识进一步了解。
- (3) 依据所整理的文献资料等进一步完善，细化研究框架和研究计划，提出详细提纲和实施方案。

第二步，课题开展阶段

(2024 年 11 月-2025 年 3 月)

- (1) 在多个社交媒体上收集社交媒体用户的个人特征、历史购买记录、互动行为（如点赞、评论、分享）、浏览历史、购买意向、购买频率和购买金额等购买行为的相关数据。
- (2) 进行数据清洗和预处理，包括去除无关信息、处理缺少值、异常值检测和处理、数据转换。
- (3) 进行特征工程，根据研究目标，从预处理后的数据中提取有用的特征。这包括用户的人口统计特征、行为模式、社交网络结构、情感分析结果等。同时，使用一些技术手段来提取文本数据特征。

(2025 年 4 月)

- (1) 构建模型。使用 RIME 算法优化 CNN-LSTM-Attention 模型的超参数，然后定义卷积层、LSTM 层和注意力机制层，以及其他可能的全连接层或池化层。最后用这些层的组合帮助模型学习数据中的复杂模式。
- (2) 模型训练。将预处理后的数据分为训练集和验证集。使用训练集对模型进行训练，同时在验证集上评估模型的性能，以避免过拟合。在训练过程中，不断优化模型。

(2025 年 7 月-2025 年 10 月)

- (1) 将经过训练和评估的模型部署到实际的应用环境中。对新数据进行预测，进而进行统计分析以得出社交媒体用户购买行为的特点。
- (2) 根据我们的研究结果，撰写相关研究论文。

第三步，成功形成阶段

(2026 年 1 月-2026 年 5 月)

- (1) 修改并最终确定论文。
- (2) 撰写研究报告，准备课题汇报与成果汇编。
- (3) 进行制定对应业务策略

6. 项目组成员分工:

李晓达：负责进行人员安排，统筹整个项目流程，填写申报书

徐悦玲：负责数学建模和相关算法的实现，辅助填写申报书

孟祥成：负责数据收集，国内外相关文献的查询与整理

王薪颖：负责查找与整理资料，整理数据，研究论文

杨文婷：负责研究论文，论文的撰写

徐嘉怡：负责实地考察，并采集汇总数据

四、预期成果

1.发表一篇“基于 RIME-CNN-LSTM-Attention 回归预测模型的平台用户购买行为研究”为主要研究内容的论文。

2.团队编写或者整理程序 RIME-CNN-LSTM-Attention 预测基于雾凇算法优化卷积神经网络结合注意力机制的长短记忆网络实现对社交媒体用户购买行为的回归预测。

3.撰写一篇“基于 RIME-CNN-LSTM-Attention 的多变量回归预测模型的社交媒体用户购买行为研究”的研究报告。

五、经费预算

1.论文发表费（计划发表 1 篇论文，需论文发表费）6000 元

2.差旅费（参加相关方面学术会议）4000 元

3.资料费（收集资料，购买相关书籍，工具书，网上有偿资料下载等）3000 元

4.打印复印费（数据打印分析，论文文献参考资料，网上资料打印复印，论文修改，表格填涂等）2000 元

5.计算费（大量数据处理需要高质量计算机进行运行，以及数据处理所需要的软件购买）2000 元

6.专家咨询费（咨询财经专家和数学专家）3000 元

总计 20000 元

六、指导教师意见

签名：

年 月 日

七、学院意见（项目所在学院）

学院负责人签名： （学院公章）

年 月 日

八、学校推荐意见

学校负责人签名： （学校公章）

年 月 日

注：表格栏高不够可增加。