

基于深度学习的物流成本预测模型研究

杨程¹, 武筱琰², 罗圆^{1*}

(1.广西计算中心有限责任公司, 广西壮族自治区南宁市, 530000;

2.桂林电子科技大学, 广西壮族自治区桂林市, 541000;

* 通讯作者, luoyuan_pro@foxmail.com)

摘要: 为提升物流成本预测精度, 本文提出一种双分支残差融合神经网络 (Dual-Branch Residual Fusion Network, DBRFN)。该模型采用上下并行双分支结构, 设计差异化激活函数, 配合残差连接强化跨层信息传递, 实现对物流成本多维度特征的精细化提取与深度表征。经特征融合实现多维度信息互补, 最终输出预测结果。实验结果表明, DBRFN表现出优异的预测性能: 决定系数 (R^2) 达0.9983, 平均绝对百分比误差 (MAPE) 为2.06%, 平均绝对误差 (MAE) 为274.21元, 预测精度显著。这种双分支残差融合设计通过特征互补机制, 不仅实现了全成本区间的精准预测, 更凭借良好的泛化能力与稳定性, 为物流成本智能化预测提供了可靠技术支撑。

关键词: 深度学习; 神经网络; 物流成本预测; 预测精度; 数据融合

引言

在交通运输网络复杂化、货运需求高涨的背景下, 物流成本的精准预测对交通规划、运力调度及企业决策意义重大 [1]。公路整车运输价格波动、铁路散货运输成本风险及冷链物流动态成本调整等场景, 均需通过科学的模型解析物流成本变化的内在机制与影响因素, 以此应对油价、供需等不确定性挑战。随着行业数字化转型, 货运数据从运量、距离等结构化信息, 扩展至实时路况、气象等多源异构数据。这既丰富了预测特征, 也对模型的特征提取、适应及泛化能力提出更高要求。

早期研究多采用统计模型方法, 如线性回归、时间序列分析与指数平滑等。例如Yousefi等人运用多元线性回归模型, 分析了加拿大货运体系中重量、距离与运量对收入的影响 [2]。这类方法虽解释性有优势, 但难以刻画交通系统中普遍存在的非线性关联。近年来, 机器学习与深度学习因在高维特征处理、非线性建模上的优势成为热点。例如, Boresta等人结合运筹学与eXtreme Gradient Boosting (XGBoost)、Light Gradient Boosting Machine (LightGBM) 构建物流服务成本预测框架 [3]; Jang等人验证了LightGBM在货运经纪平台的最优预测性能 [4]。在深度学习领域, 图卷积网络Graph Convolutional Network (GCN) 为揭示多节点物流网络成本传导关系提供思路 [5]; 卷积与编码器—解码器结构实现了细分货运流量的精细化预测 [6]。

尽管深度学习在物流成本建模中潜力显著, 但在多源异构数据动态融合、时空依赖关系与长时序特征捕捉上仍有优化空间。基于此, 本文聚焦神经网络, 设计适应性与表达能力更强的网络架构, 以提升物流成本预测精度与鲁棒性, 为复杂场景智能化预测提供技术支撑。

1. 数据预处理

本研究所采用的数据集来源于某大型港口的货运订单管理系统, 涵盖过去一年间经该港口流转的陆运及内河航运货物交易记录。为提升数据的代表性, 对原始订单数据进行初步筛选, 剔除存在异常波动的极端运费值, 仅保留运费处于2,000元至60,000元区间的有效订单。

原始数据集中共包含34个特征, 涵盖货物属性、运输路线以及运输时间等多个维度。经筛选, 选取其中15个与运费高度相关的核心变量用于后续建模。剔除含“dull”标记或空白字段的异常行, 删除重量、费率、里程、周转量为零的无效订单数据, 以确保数据质量。针对类别型变量, 统一进行数值化处理。通过定义映射字典, 对出发港口和目的港口进行合并编码, 为每个港口分配唯一整数标识。对“货物类型”列提取其中的数字部分进行标注化处理。对于“货物名称”、“运输路线”等类别特征分别构建值-编码映射关系, 并生成编码对照表。为进一步提升数据质量, 引入基于密度的聚类算法DBSCAN (Density-Based Spatial Clustering

of Applications with Noise) 识别并剔除离群样本。设置参数Epsilon=0.5, Minimum Samples=3, 实现对低密度区域样本的自动剔除。

在完成上述处理后, 数据集中共有261574条数据。对数据集中的特征变量与目标变量进行分离, 按8:2的比例划分训练集和测试集, 同时设定随机种子为42, 以保证数据划分的稳定性和实验的可重复性, 从而为后续的模型训练和验证提供可靠的数据基础。

2. 物流成本预测模型

2.1. 模型整体架构

针对物流成本预测问题的复杂性和多维度特性, 本研究提出一种双分支残差融合神经网络模型 (DBRFN)。模型采用上下并行的双分支结构 (图1), 通过差异化的激活函数设计和残差连接机制, 旨在实现物流成本多维度特征的精细化提取。

模型的输入层接收经过预处理的多维物流特征向量, 其中包含运输距离、货物类型等15个特征向量。上分支 (Branch A) 采用Swish激活函数, 专注于挖掘数据中的复杂非线性关系; 下分支 (Branch B) 结合ReLU和tanh激活函数, 着重解析数据的稳定特征模式。两条分支均引入残差连接, 以增强信息传递能力。在特征整合阶段, 双分支的输出通过拼接形成融合特征向量, 最终送入线性激活函数 (linear) 层, 线性激活函数将融合特征向量直接映射为具体的物流成本预测值, 完成从特征空间到预测结果的最终转化。

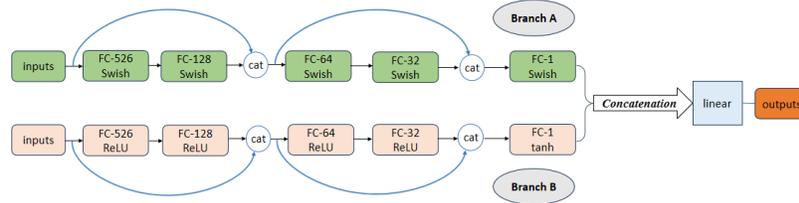


图1 DBRFN神经网络图

图注: 全连接层 (FC) 后跟随Batch Normalization和Dropout (0.3), 分支输出层及残差拼接操作除外。

2.2. 双分支特征提取机制

上分支以Swish函数为核心激活函数, 利用其优越的非线性变换能力和梯度传播特性, 捕捉物流成本数据中的复杂非线性关联和动态变化趋势。分支整体包含5个全连接层, 神经元数量依次为526、128、64、32和1, 实现从高维特征到抽象特征的层级提取。分支设计了两处残差连接: 一是将原始输入特征与第二层输出特征拼接, 保留基础特征信息以避免底层信息丢失; 二是将第四层输出特征与第二层输出特征拼接, 强化跨层特征复用, 提升多尺度特征的协同表达。同时, 全连接层之间均加入批量归一化 (Batch Normalization, BN) 和随机失活 (Dropout), 其中Dropout的失活概率设为0.3。通过规范化特征分布和随机失活神经元, 抑制模型过拟合, 提升模型在未知数据上的泛化能力。

下分支主要利用ReLU激活函数提取数据中的关键稳定特征, 并在输出层引入tanh函数, 将特征映射至[-1, 1]区间, 强化边界特征的区分度。该分支同样包含5个全连接层, 神经元数量与上分支保持一致。类似上分支, 下分支也设置了两处残差连接: 通过拼接原始输入与第二层输出, 保留基础特征的完整性; 通过拼接第二层输出与第四层输出, 实现跨层特征互补, 增强模型对特征层级关系的学习能力。同样, 每两层全连接层间之间配置BN和Dropout (0.3), 与上分支形成统一的正则化策略, 保障模型训练的稳定性与特征提取的可靠性。

2.3. 特征融合与输出层设计

为充分整合双分支提取的差异化特征, 模型采用特征拼接 (Concatenate) 策略进行融合: 将上、下分支输出的特征进行拼接, 形成包含多尺度、多类型信息的综合特征向量。融合后的综合特征向量通过一个全连接层进行最终信息整合, 该层采用线性激活函数, 将高维特征映射为连续的预测值, 满足物流成本预测的回归任务需求, 最终生成单一预测值作为模型的输出结果。

2.4. 模型正则化与优化策略

为提升模型的泛化能力并确保训练稳定性，模型采用了多层次正则化策略，包括BN和Dropout。这些措施帮助加速收敛、缓解梯度消失问题，并有效抑制过拟合。优化方面，模型选用Adam优化器，其通过自适应调整学习率（初始学习率设为0.001），结合一阶矩估计（ $\beta_1 = 0.9$ ）和二阶矩估计（ $\beta_2 = 0.999$ ）动态更新参数，兼顾收敛速度与稳定性。损失函数采用均方误差（MSE），以最小化预测值与真实值的平方差为优化目标，并将平均绝对误差（MAE）作为辅助评估指标。模型训练时，设置批量大小为2000，总训练轮次为100，同时划分20%的训练数据作为验证集，以实时监控模型在未参与训练数据上的表现，从而确保模型在复杂物流成本数据上的拟合效果与泛化能力达到最佳平衡。

3. 实验及结果

3.1. 对比实验

本研究设计了对比实验，将完整模型与组成它的两支单一分支网络进行对比分析。实验结果显示，DBRFN在各评价指标中均表现最优：决定系数（ R^2 ）在epoch 100时达到0.9983，高于单一分支（Branch A为0.7952，Branch B为0.8452（如图2），表明其对数据规律的拟合能力更强；平均绝对百分比误差（MAPE）在epoch 100时低至2.06%，优于Branch A的51.93%与Branch B的16.58%（如图3）；平均绝对误差（MAE）与均方根误差（RMSE）同样呈现最优表现，在epoch 100时MAE为274.21元（如图4），RMSE为519.9元（如图5），均显著低于单一分支，充分验证了该模型在物流成本预测精度上的优势。

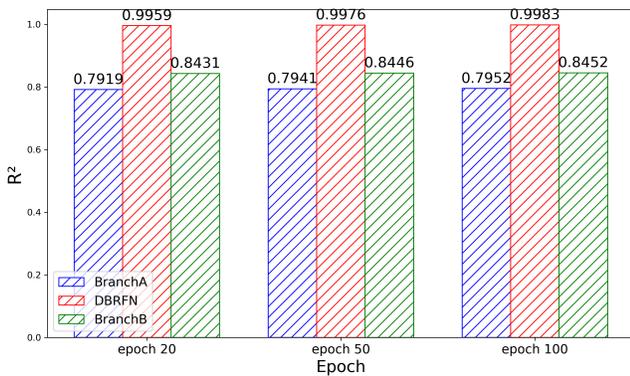
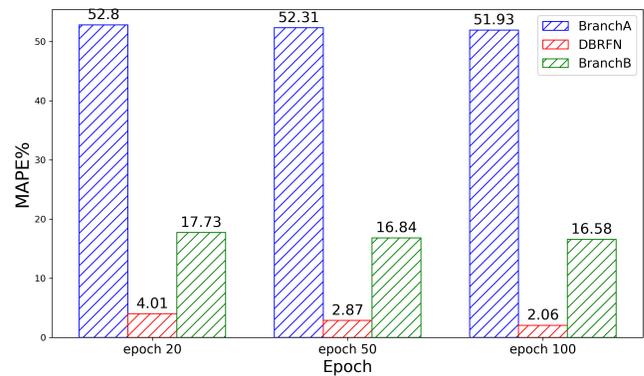
图2 R^2 图

图3 MAPE图

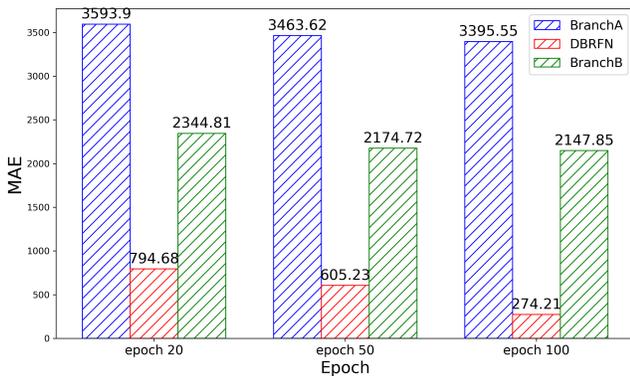


图4 MAE图

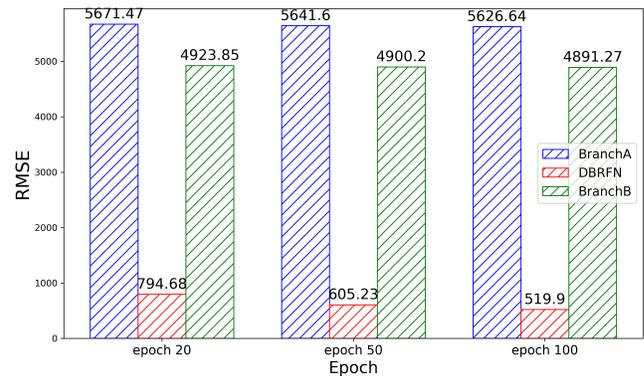


图5 RMSE图

3.2. 独立分支网络的预测偏好差异分析

为验证两支独立神经网络的预测偏好差异，对拆分后的Branch A和Branch B分别进行独立训练，沿用与DBRFN相同的训练参数：batch size=2000，epochs=100，Adam优化器及MSE损失函数），并在相同测试集上

评估其预测性能。通过将两支网络的预测值与真实物流成本值进行比对，结合折线图中预测曲线与真实曲线的贴合度分析，发现：**Branch A**在较高物流成本数据的预测中表现出优势，能够捕捉高成本数据的波动特征与潜在规律，对峰值成本的预测贴合度较好（图6）；**Branch B**则更适配较低物流成本数据的预测，在处理成本相对平稳、数值偏低的样本时，可精准拟合低成本区间的变化趋势（图7）。

当将这两支适配不同成本区间的神经网络进行组合，构建成完整的组合神经网络后，预测效果得到显著提升。**DBFRN**网络充分整合了两支单网络的优势，对于物流成本数据中的高成本峰值、低成本平稳段，均能实现更精准的预测。从**DBFRN**对应的折线图可见（图8），预测值（红线）与真实值（蓝线）贴合度大幅提升，无论高低成本区间，波动趋势均可较好跟随，有效弥补了单支网络的局限性，展现出对全成本区间良好的预测能力。

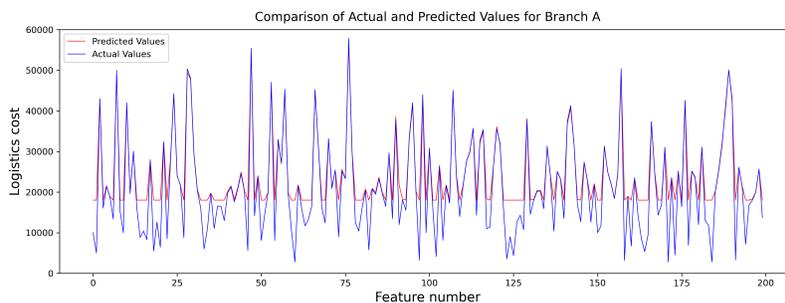


图6 Branch A物流成本实际值与预测值对比图

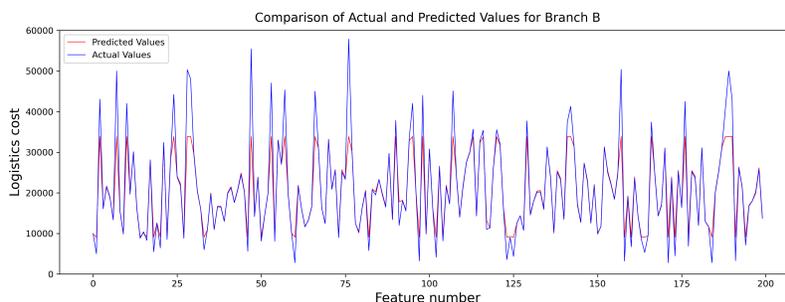


图7 Branch B物流成本实际值与预测值对比图

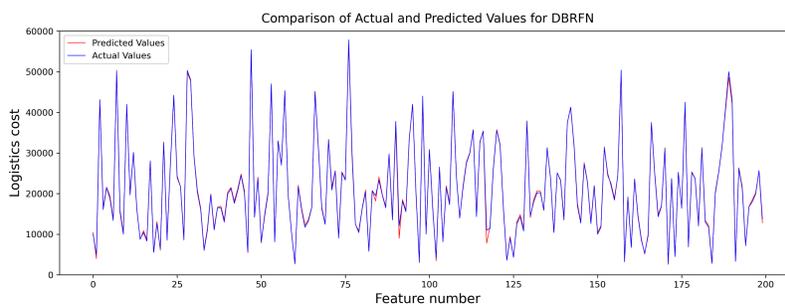


图8 DBFRN物流成本实际值与预测值对比图

3.3. 不同模型训练损失的收敛过程对比

为直观呈现不同模型的训练收敛过程，绘制了**DBFRN**、**Branch A**与**Branch B**的训练损失（Loss）随迭代轮次（Epoch）变化的对比曲线（如图9）。从图中可知，训练初期（Epoch < 10），各模型损失均快速下降，**DBFRN**凭借融合结构优势，损失衰减速率显著高于单一分支；随着训练推进（10 ≤ Epoch ≤ 30），**DBFRN**损失持续下降，**Branch A**、**Branch B**的损失下降则渐趋平缓并进入稳定阶段；当Epoch ≥ 30后，**DBFRN**损失已降至极低水平并保持稳定，而**Branch A**、**Branch B**的损失分别维持在0.21、0.15左右。整体来看，**DBFRN**的损失曲

线始终处于最低位，且更早进入收敛状态，有力印证了其在训练效率与拟合能力上相较于单一分支网络的显著优势。

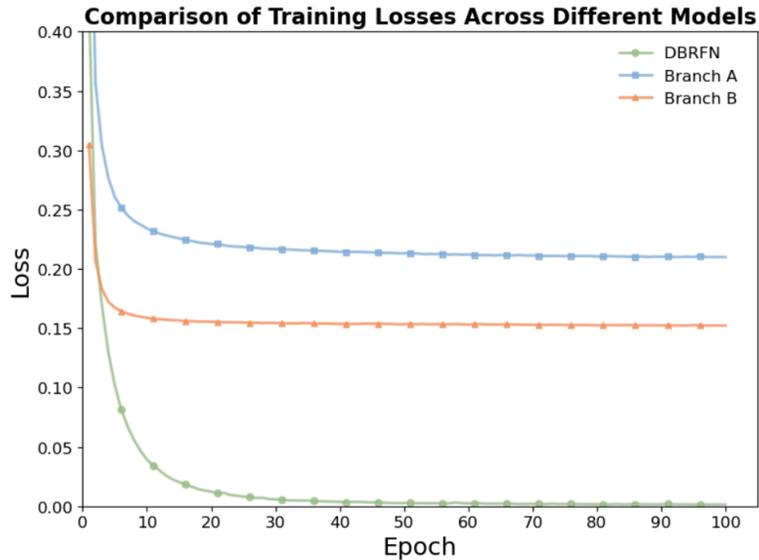


图9 模型损失函数变化曲线图 R^2

4. 总结

本文围绕物流成本预测问题，提出了一种创新的双分支残差融合神经网络（DBRFN），有效融合了Swish、ReLU与tanh激活机制，结合残差连接与特征拼接策略，实现了对物流多维度特征的深层建模与融合表达。该模型在训练过程中表现出良好的收敛性与稳定性，具备较强的适应性与泛化能力，满足全成本区间的预测需求。本研究为物流成本的智能化预测提供了高效可靠的技术路径，同时也为复杂数据背景下的深度学习模型设计提供了理论参考。

参考文献

- [1] 徐志刚, 申丹丹, 高赢, 等. 基于文献计量的综合交通研究综述 [J]. 交通运输工程学报, 2025, 25(2): 37-60.
- [2] YOUSEFI J, ASHTAB S, YASAEI A, et al. Multiple linear regression analysis of canada's freight transportation framework [J]. Logistics, 2023, 7(2): 29.
- [3] BORESTA M, PINTO D M, STECCA G. Bridging operations research and machine learning for service cost prediction in logistics and service industries [J]. Annals of Operations Research, 2024, 342(1): 113-139.
- [4] JANG H S, CHANG T W, KIM S H. Prediction of shipping cost on freight brokerage platform using machine learning [J]. Sustainability, 2023, 15(2): 1122.
- [5] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks [J]. arXiv preprint arXiv: 1609.02907, 2016.
- [6] LLORET I, TROYANO J A, ENRIQUEZ F, et al. Two deep learning approaches to forecasting disaggregated freight flows: convolutional and encoder-decoder recurrent [J]. Soft Computing, 2021, 25(12): 7769-7784.