

基于PN和CNN-LSTM-ATT的航班延误预测

吴 涔¹, 叶 宁^{1,2}, 王 甦¹, 季翔宇¹

(1. 南京邮电大学 计算机学院、软件学院、网络空间安全学院, 江苏 南京 210023;

2. 江苏省无线传感网高技术研究重点实验室, 江苏 南京 210023)

摘 要: 航班延误预测对提高机场地面保障效率具有重要参考意义。针对目前航班地面保障流程复杂多变以及航班过站延误预测精度不高的问题, 提出了一种基于Petri Net和融合预测模型CNN-LSTM-ATT的航班延误预测模型。首先, 根据机场航班实际地面保障流程抽象构建离港航班地面保障作业Petri Net模型, 获取保障流程中的关键作业时长成为动态特征; 其次, 将动态特征、航班信息、延误信息和天气信息输入CNN-LSTM-ATT模型中进行特征提取和分类预测, 模型中引入注意力机制, 通过注意力权重突出关键数据信息的影响, 进一步挖掘重要特征之间的内部规律。实验结果显示, 该融合模型准确率相比独立模型提升了6%, 达到98.1%。通过对不同模型的对比表明该模型能较好地应对场面流程变化并且具备较好的延误预测能力。

关键词: 航班地面保障; 保障流程分析; 延误预测; Petri Net; CNN-LSTM-ATT

中图分类号: TP183

文献标识码: A

文章编号: 1673-629X(2023)04-0213-08

doi: 10.3969/j.issn.1673-629X.2023.04.031

Flight Delay Prediction Based on Petri Net and CNN-LSTM-ATT

WU Cen¹, YE Ning^{1,2}, WANG Su¹, JI Xiang-yu¹

(1. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China;

2. Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing 210023, China)

Abstract: Flight delay prediction is an important reference to improve the efficiency of airport ground service. Nowadays, the flight ground service process is complex and changeable, and the prediction accuracy of flight transit delay is not high, thus a flight delay prediction model based on Petri net and fusion prediction model CNN-LSTM-ATT is designed. Firstly, the Petri Net model of the departing flights ground service process was abstractly constructed according to the actual ground service process of flights, and the critical service operation duration became the dynamic feature. Secondly, the dynamic feature, flight information, delay information and weather information are input into the CNN-LSTM-ATT model for feature extraction and classification prediction. The attention mechanism is introduced to highlight the influence of key data through attention weight and further mine the internal laws of important features. The results show that the fusion model improves the accuracy by 6% compared with the independent model, and the final accuracy of the experiment reaches 98.1%. The comparison of different models shows that the model can better cope with the changes of scene and has better delay prediction ability.

Key words: flight ground service; service process analysis; delay prediction; Petri Net; CNN-LSTM-ATT

0 引言

随着国民经济和民航业的快速发展, 航班延误产生的压力也是越来越大。航班延误不仅会影响旅客的出行计划, 给旅客带来时间和经济损失, 同时也增加了航空公司的运营成本, 影响了机场和空管部门的正常运营。然而航班的延误通常是由多方面因素造成的, 因此延误是难以完全避免的^[1]。针对航班延误问题,

国内外学者已展开大量研究, 大多采用支持向量机、决策树、神经网络等传统的机器学习算法。为了提高机场地面业务的处理效率, Patrick等^[2]提出一种多智能体计划方法重组机场地面业务管理, 减少因各业务部门间信息共享延迟而造成的延误, 但仅停留在理论层面, 仍需要更多实际场景进行验证。杨珏^[3]通过不同启发算法的遗传算法对保障服务无延误且调度成本最

收稿日期: 2022-05-12

修回日期: 2022-09-15

基金项目: 江苏省科技重点研发计划(社会发展)(BE2020713)

作者简介: 吴 涔(1998-), 女, 硕士研究生, 研究方向为深度学习; 叶 宁(1971-), 女, 教授, CCF高级会员(10059S), 研究方向为无线传感网络信息与安全。

低的车辆路径模型进行求解,找到最佳车辆调度路径从而提高地面保障效率。Wu 等^[4]从机场、航空公司和航班三个方面更为全面地考虑影响航班延误的因素,为降低模型复杂度,使用主成分分析法进行降维,将历史数据作为先验知识输入到 SVM 模型,从而预测延误时间。Jiang 等^[5]提出了一种基于长短期记忆网络(Long Short-Term Memory, LSTM)的模型,分析了时间特征、航班属性、机场和城市天气情况等多维度因素,结合过去一段时间的延误情况,同时预测了航班延误率、平均延误时间等多个延误指标,该方法有助于提前发现航班延误趋势的变化,为机场调度提供决策。但上述研究均只采用了单一模型,很难挖掘不连续数据之间的潜在信息。因此,有研究^[6-10]将 CNN 和 LSTM 模型融合进行预测和分类,但对于模型来说,每条数据的重要性是不一样的,出现延误情况的数据项含有更多的隐含信息,这种重要性差异是很难在 CNN 模型中体现的,使得模型难以正确评估样本的重要性。通过模仿人的注意力, Vaswani^[11]提出了注意力机制(Attention Mechanism)的概念,使得神经网络模型能关注更加有效的信息。

大型机场地面保障环境复杂多变,对于外界因素的改变,人工和机器学习模型都难以及时响应。因此,为了提高预测准确率,对保障流程系统进行分析具有重要的现实意义。Petri Net 对于具有异步、并发和冲突等关系的系统具有良好的建模能力,因此被广泛应用于 workflow 管理、故障诊断和电力系统等领域。Zhu^[12]使用 Petri Net 描述机场停机坪的交通动态,利用着色时序 Petri 网对各运行区域进行动态建模,反映停机坪交通系统的控制和资源占用规则,提出了停机坪冲突控制机制。Davidrajuh 等^[13]提出了一种通过扩展聚类算法来自动检测关键元素的算法,将大规模离散事件构成的 Petri Net 模块化,降低了模型的复杂度,便于系统的进一步分析。Luo 等^[14]将排队论和 Petri Net 相结合来实现服务系统态势感知,该模型对通用服务系统具有重要意义,能够更加全面地描述一般服务系统,并从宏观上展现系统之间的联系。

因此,为了提升航班延误预测模型的灵活性和准确性,及时分析和应对外部条件变化,该文提出了一种将时序 Petri Net 与融合模型 CNN-LSTM-ATT 结合构建航班延误预测模型的研究方法,根据机场实际航班地面保障流程,构建航班保障流程 Petri Net,通过矩阵推理对保障流程的关键路径进行计算,分析其中关键流程作业成为动态特征。然后,将航班信息、机场延误信息、机场天气信息以及关键流程信息输入到融合预测模型进行延误分类预测,通过国内大型机场实际数据对模型的有效性进行分析验证。

1 流程描述

1.1 航班过站保障流程

航班过站保障流程是指过站航空器从上轮挡开始到撤轮挡完成期间内做出的一系列保障作业的集合,根据《航班安全运行保障标准》中的规定,不同机型具有不同的保障工作时长规定,该文主要研究当前大型机场中最常见的 A320 系列、B737 系列等型号航空器。航班地面保障工作流程中包含了许多并联工作关系和串联工作关系,各项作业之间存在一定的约束关系。保障流程按照服务类型可以分为航空器服务、旅客服务以及货舱服务。

考虑到旅客体验以及安全问题,保障流程中的部分工作之间存在一定的先后关系和逻辑次序,例如,在客舱清洁、配餐食和加油工作完成后才能安排旅客登机工作,货舱卸舱完成后才可以安排货物装舱等,所以在工作流中一项作业的超时有可能会造成后继作业的延期开启。该文研究的保障作业流程如图 1 所示。

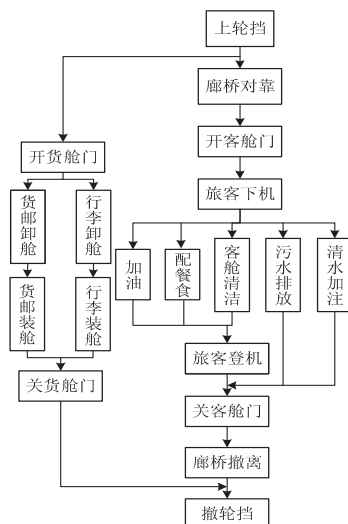


图 1 航班保障作业流程

1.2 航班延误预测流程

机场地面环境复杂,航班在进行保障的过程中会受到很多外界因素的干扰,过站保障时长会随之产生波动,现有的神经网络预测具有较高的准确率,但缺乏对保障工作运作的实时分析,因此目前在很多大型机场中,为保障航班准时出发的保障作业时长调整往往在异常发生后才能进行。对于机场外部条件的变化,现有的预测模型是很难及时应对的。

经分析,航班在出现延误时,地面部分保障作业会出现超出规定时长的情况,为分析具体保障作业对整体航班过站时长的影响,因此研究设计首先参考机场实际航班保障作业流程图,抽象化建立地面保障流程 Petri Net 模型,计算历史平均保障时间来获取动态关键保障作业,将其作为动态特征与航班基本特征、场面延误信息以及外部天气信息进行数据预处理后输入到

融合预测模型。模型使用航班宏观特征以及微观特征进行延误分类来解决预测模型缺乏灵活性的问题。

2 时序 Petri Net 建模与分析

2.1 Petri Net

Petri Net (PN) 是一种适用于描述离散事件系统并发、冲突和死锁等动态行为的图形化的数学工具^[15]。航班保障过程包含了一系列具有先后次序和逻辑次序的作业。PN 由库所 (Place)、变迁 (Transition) 和有向弧 (Directed arcs) 连接而成,库所表示作业状态,变迁表示作业变化过程,有向弧用来表示库所和变迁之间的依赖关系。

PN 模型可表示为: $PN = (P, T, F)$, 其中, $P = \{p_1, p_2, \dots, p_m\}$ 是库所集合, 每一个 p_i 表示保障作业任务状态, $T = \{t_1, t_2, \dots, t_n\}$ 是变迁集合, 每一个 t_i 表示保障作业过程, 且 $P \cap T \neq \emptyset, P \cup T \neq \emptyset, F \subseteq (P \times T) \cup (T \times P)$ 表示的是 T 和 P 之间的有向弧集合。

$PN = (P, T, F)$ 为最基本的网络结构, PN 会通过托肯 (Token) 来反映流程进行到何种状态, 托肯常常用来表示资源的数量。PN 系统行为的本质是资源的流动, 保障过程中流动的可以理解为不断改变工作状态标签的保障人员。保障过程中局部作业的延期会波及后续作业, 从而导致航班延误。PN 模型聚焦保障作业节点, 以便于预测模型对保障流程内部作业的关注。例如, t_1 是关舱门廊桥撤离, t_2 是撤轮挡作业, 两项作业具有严格的先后次序, 只有前序变迁完成时托肯流入后续库所, 托肯的流动是矩阵推理的基础。PN = (T, P, F, M) 表示的是一个带有标识的 Petri Net,

$M(P)$ 可以用来表示库所中的托肯, 因此, PN 的标识可以通过 m 维正整数向量来表示, 从而对流程进行分析计算。

2.2 航班保障流程关键特征提取

以国内某大型机场航空器地面运行保障全流程为例, 其离港航班地面保障作业如表 1 所示。总共包含 15 项保障作业和 1 项虚拟作业, 相关作业具有一定的串行或者并行的逻辑关系。

表 1 保障作业与变迁

保障作业	对应 Petri Net 变迁
上轮挡作业	t_1
廊桥对靠 开舱门	t_2
旅客下机	t_3
配餐食作业	t_4
污水排放作业	t_5
清水加注作业	t_6
客舱清洁作业	t_7
加油作业	t_8
旅客登机	t_9
货舱卸行李	t_{10}
货舱装行李	t_{11}
货舱卸货	t_{12}
货舱装货	t_{13}
关舱门 廊桥撤离	t_{14}
撤轮挡作业	t_{15}
虚拟作业	t_{16}

根据上述作业对应变迁以及作业之间的紧前后继关系, 对 1.1 节的航班保障作业流程抽象构建离港航班地面保障作业 Petri Net 模型, 如图 2 所示。

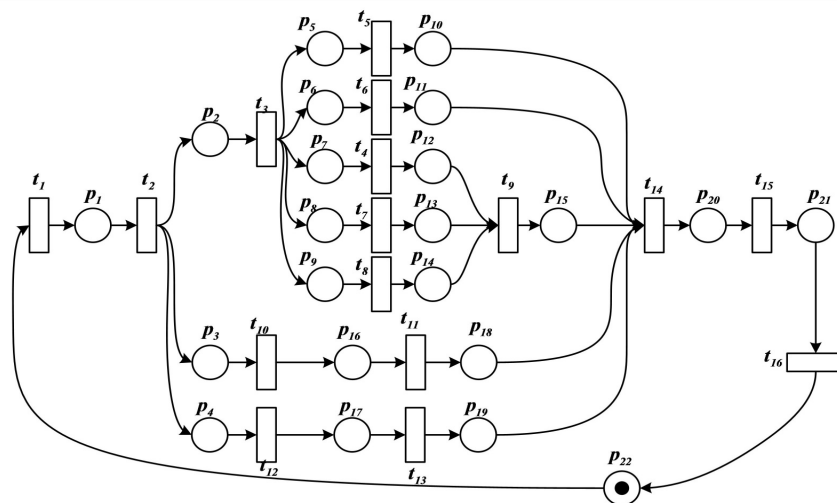


图 2 地面保障作业 Petri Net 模型

为了获得该模型的关键路径, 获取其中关键流程作业用于后续分析, 需要计算获得 PN 中 S 不变量来对其进行分析, PN 的关联矩阵可以用来表示其中的状态并用于描述状态转换。

定义 1: 设 $PN = (P, T, F, M_0)$ 为一个时序 Petri Net 模型, 库所集合 $P = \{p_1, p_2, \dots, p_m\}$, 变迁集合 $T = \{t_1, t_2, \dots, t_n\}$, 该系统中 (P, T, F) 可以用 $n \times m$ 的矩阵表示, $C = [c_{ij}]_{n \times m}$, 称 C 为该网络系统的关联矩阵。

根据定义 1 可知,关联矩阵 C 中一行元素个数对应的是系统中的变迁个数,每个元素用来表示托肯数的改变。

定义 2: 设 $PN=(P,T,F,M_0)$ 为一个时序 Petri Net 模型,当且仅当 m 维列向量 Y 满足以下条件:

$$(1) C_{n \times m} \cdot Y_{m \times 1} = 0;$$

(2) 向量 Y 中的任一元素均为整数。则称 Y 是 PN 的 S 不变量。

针对图 2 建立的地面保障作业 PN 模型建立关联矩阵,获得输出关联矩阵 c_{ij}^- 和输入关联矩阵 c_{ij}^+ ,依据 $c_{ij} = c_{ij}^+ - c_{ij}^-$,获取模型的关联矩阵,如式(1)所示。

$$C \cdot Y = 0 \quad (1)$$

式中, C 为模型的关联矩阵, Y 为需要求解的 S 不变量,可以获得模型的所有 S 不变量,该文建立的 PN 模型通过矩阵推理可获得 7 个一维矩阵,如式(2)所示。

$$Y_1^T = (1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1)$$

$$Y_2^T = (1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1)$$

$$Y_3^T = (1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1)$$

$$Y_4^T = (1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1)$$

$$Y_5^T = (1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1)$$

$$Y_6^T = (1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1)$$

$$Y_7^T = (1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1) \quad (2)$$

在求解获得的不变量中,数字‘1’代表该库所畅通,根据计算获得的 7 个 S 不变量,可以获得 7 条保障作业路径,因此求得的不变量如表 2 所示。

表 2 S 不变量及对应路径

变量名	路径	变迁集合
Y_1	$P_1 P_2 P_5 P_{10} P_{20} P_{21} P_{22}$	$t_1 t_2 t_3 t_5 t_{14} t_{15} t_{16}$
Y_2	$P_1 P_2 P_6 P_{11} P_{20} P_{21} P_{22}$	$t_1 t_2 t_3 t_6 t_{14} t_{15} t_{16}$
Y_3	$P_1 P_2 P_7 P_{12} P_{15} P_{20} P_{21} P_{22}$	$t_1 t_2 t_3 t_4 t_9 t_{14} t_{15} t_{16}$
Y_4	$P_1 P_2 P_8 P_{13} P_{15} P_{20} P_{21} P_{22}$	$t_1 t_2 t_3 t_7 t_9 t_{14} t_{15} t_{16}$
Y_5	$P_1 P_2 P_9 P_{14} P_{15} P_{20} P_{21} P_{22}$	$t_1 t_2 t_3 t_8 t_9 t_{14} t_{15} t_{16}$
Y_6	$P_1 P_3 P_{16} P_{18} P_{20} P_{21} P_{22}$	$t_1 t_2 t_{10} t_{11} t_{14} t_{15} t_{16}$
Y_7	$P_1 P_4 P_{17} P_{19} P_{20} P_{21} P_{22}$	$t_1 t_2 t_{12} t_{13} t_{14} t_{15} t_{16}$

根据航空公司给出的历史航班地面保障作业时间数据输入到 PN 模型,计算上述各路径的总作业时长,对比获得关键路径,计算获取关键路径中时间占比最大的作业和在出现延误情况下时间波动最大的作业,具体数据计算方法见 4.1 节。保障流程中的关键作业项会受到天气信息、航班流量和机位调整等因素出现

改变,定期将历史数据输入模型中获取到的关键作业项会出现变化,动态变化的节点项成为动态微观特征,再融合其他常规特征项输入到融合预测模型。

3 基于 CNN-LSTM-ATT 的延误预测模型

由于航班过站期间场面复杂多变,受多种因素影响,单一预测模型不够稳定,容易忽略重要特征信息,该文融合多种模型,利用注意力机制赋予重要特征更多的权重,以此提高模型预测精度。

3.1 CNN 特征提取

卷积神经网络 (Convolutional Neural Networks, CNN) 的主要功能是从数据中提取特征。CNN 由卷积层、池化层和全连接层组成^[16]。卷积层是提取特征的关键,通过卷积核对输入数据进行特征提取,卷积核数量增加会使得提取出的特征更为抽象,通过 RELU 激活函数加入非线性因素解决问题,该文采用最大池化层忽略一些不重要的信息加快计算速度以及防止过拟合。全连接层可以将池化层获得的数据扩展成一维向量,方便后续输出操作。但 CNN 再进行预测时很难学习到序列数据之间的关系,航班数据之间具有依赖关系,所以需要将 CNN 与 RNN 进行结合。

3.2 LSTM 基本原理

LSTM 是一种改进的循环神经网络,它解决了模型训练中梯度消失的问题并且能够学习序列信息的依赖关系^[17],现如今 LSTM 由于其良好的学习性能,已经被应用于很多场景中。

LSTM 的基本神经单元包含了遗忘门、输入门和输出门,门可以理解为对数据的一次处理。遗忘门主要是过滤对模型训练无用的干扰信息,根据上一时刻的输出 h_{t-1} 和当前输入 x_t 共同产生一个值来决定是否让上一时刻学到的信息通过该单元;输入门则会产生需要的新信息,使用 sigmoid 选出的信息和 tanh 来生成新的候选信息结合进行信息更新;最后输出门控制模型的输出,LSTM 当前状态单元的输出是由前一时刻的数据和当前输入的数据共同生成的,具体计算公式如式(3)所示。

$$\begin{aligned}
 f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f) \\
 i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i) \\
 O_t &= \sigma(W_o[h_{t-1}, x_t] + b_o) \\
 g_t &= \tanh(W_g[h_{t-1}, x_t] + b_g) \\
 C_t &= g_t \cdot i_t + f_t \cdot C_{t-1} \\
 h_t &= O_t \cdot \tanh(C_t)
 \end{aligned} \quad (3)$$

式中, f_t 、 i_t 、 O_t 分别代表遗忘门、输入门和输出门, C_t 表示当前单元状态, h_t 表示隐层状态, W 和 b 分别表示权重和偏移量, σ 表示 sigmoid 激活函数。

在改进的 LSTM 中,模型训练过程中会根据设置

的概率临时随机删除隐藏层中的神经元,来防止过拟合,未删除的神经元进行连接,输入信息前向传播,获得的损失值进行反向传播,在模型结束训练后更新未被删除的神经元的权值,临时删除的神经元也要重新恢复,重复这个过程直到数值收敛。

加入 dropout 的 LSTM 内部计算式如式(4)所示。

$$\begin{aligned} \tilde{\mathbf{y}}^{(l)} &= \mathbf{r}^l \cdot \mathbf{y}^l \\ \mathbf{z}_i^{(l+1)} &= \mathbf{W}_i^{(l+1)} \tilde{\mathbf{y}}^{(l)} + \mathbf{b}_i^{(l+1)} \\ \mathbf{y}_i^{(l+1)} &= f(\mathbf{z}_i^{(l+1)}) \end{aligned} \quad (4)$$

式中, \mathbf{W} 和 \mathbf{b} 是对应神经元的权重和偏移量, \mathbf{r} 是由伯努利函数获取的随机概率,取值在 0 到 1 之间,通过概率值 \mathbf{r} 临时删除部分神经元后剩下的网络单元用 $\tilde{\mathbf{y}}^{(l)}$ 表示, $\mathbf{z}_i^{(l+1)}$ 表示的是神经单元下一时刻的输出,再经由激活函数获得输出 $\mathbf{y}_i^{(l+1)}$ 。

3.3 Attention 机制

LSTM 具有提取长序列时序特征的能力,但序列长度过长时,LSTM 的性能会下降。对于航班延误预测,只有选取足够长的序列长度,才能包含更多对过站时长存在影响的因素。注意力机制用于计算每一个前序隐层输出 $\{h_1, h_2, \dots, h_{t-1}\}$ 与当前输出 h_t 的相关性,从而构建当前时刻的输出 o_t 。

注意力机制是一种模拟人脑注意力的资源配置机制。在某一时刻人的大脑会将注意力集中在需要关注的区域,减少甚至忽略对其他区域的关注,从而获取更多值得关注的细节,减少无用信息的影响并且放大了有用的信息。注意机制通过权重分配让模型对关键信息给予足够的关注,突出重要信息的影响,从而提高模型的准确性,注意机制的引入可以有效改善 LSTM 由于序列过长而造成的信息丢失问题。LSTM 层的输出是注意力层的输入,可以采用概率分配权重的方式对权重矩阵进行计算,概率分布计算如式(5)所示。

$$\mathbf{v}_i = \text{utanh}(\mathbf{w} \mathbf{h}_i + \mathbf{b}) \quad (5)$$

按照概率分布计算注意力权重 α_i 以及输出结果 o_t ,计算方法如式(6)和式(7)所示。

$$\alpha_i = \frac{\exp(\mathbf{v}_i)}{\sum_{i=1}^m \exp(\mathbf{v}_i)} \quad (6)$$

$$\mathbf{o}_t = \sum_{i=1}^m \alpha_i \mathbf{h}_i \quad (7)$$

3.4 CNN-LSTM-ATT 模型

关键路径提取的特征、航班信息、天气信息和延误信息实际上都是独立特征,该文使用词向量表示形式,将一个时刻的特征信息与标签值用向量形式表示,从而形成新的序列数据。根据数据量确认了滑动窗口的大小为 20,步长为 1 来形成特征矩阵图,输入特征图

按照时间进行排列。

将设计好的输入数据特征输入到融合预测模型,设计的融合预测模型流程如图 3 所示。

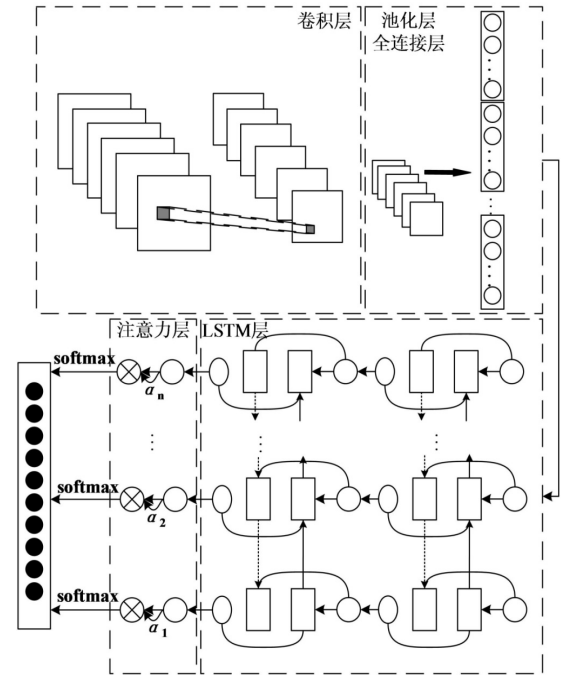


图3 CNN-LSTM-ATT 网络结构

在模型的参数设置方面,CNN 层卷积核个数为 64,卷积核的长度为 2。理论上,LSTM 隐藏层越多,预测效果越好。但是,随着层数的增加,训练时间也会大大增加,所以文中 LSTM 设置为两层,第一层 128 个隐层通道,第二层 64 个隐层通道。注意力层对输出有更大的影响,注意力层参数通过学习特征权值和分配特征输入向量实现。

4 实验与分析

4.1 实验环境以及数据概述

为保证深度学习实验能够正常进行,实验设备配备 AMD Ryzen 5 5600H with Radeon Graphics,操作系统为 Windows 10,内存 16 GB,实验使用的 Keras 的版本为 2.0.8。

实验室用的数据集来源于国内某大型机场提供的 2021 年 1 月至 5 月航班过站数据,经初步筛选,该数据集总共有 38 000 条数据,其中包括了计划出发、上轮挡时间、清洁时长、加油时长和装舱时长等多项航班过站信息。实验中使用到的气象数据来源于国家气象科学数据中心提供的地面逐小时气象数据。

此次实验模型输入特征包括四大类,具体说明如下:

第一类是动态关键路径关键作业时长,根据 3.2 节关键路径的获取方法,以 1~5 月航班历史平均作业时间为例,各项地面平均作业时间如表 3 所示。

表 3 各项作业地面历史平均用时

作业	用时/分	作业	用时/分
t_1	2.00	t_9	17.95
t_2	2.00	t_{10}	5.58
t_3	8.22	t_{11}	6.88
t_4	11.22	t_{12}	5.80
t_5	4.57	t_{13}	30.00
t_6	4.40	t_{14}	2.00
t_7	9.65	t_{15}	6.00
t_8	9.83		

根据表 3 所给出的各项作业平均用时,计算各路径的总时长。根据计算可知,2021 年上半年对航班过站时间影响最大的路径为 Y_3 , 平均路径总时长为 49.36 分钟,根据对该路径的分析将配餐时长以及旅客登机时长这两项作为该数据集的动态微观特征。

第二类是航班信息,主要包括航班号、航线、航班停机位、航班计划起飞日期、计划起飞时间、星期数、上轮挡时间、目的城市 ID、值机人数和出港行李件数等 20 项特征。

第三类是机场延误信息,包括前一天延误架次、前一小时延误架次、前一小时平均延误时间共 3 项延误信息特征。

第四类是天气信息特征,主要包括机场的航班计划起飞时间的天气现象、可见度、风力、风向、温度、大气压和目的机场天气状况共 7 项特征。

根据不同特征的属性,需要对数据进行预处理,使其更适合作为模型的输入。首先,需要对缺失数据以及具有明显逻辑错误的数据进行清洗;其次,如果直接将无序离散数据输入到模型中会让模型训练结果出现

较大的偏差,所以采用了将离散数据按出现的频次对其进行编码,连续数据按照 min-max 归一化将数据映射到 $[0,1]$;最后,按照 3.4 节中对数据输入格式的介绍,将数据转化成多个 20×32 的特征矩阵图。

4.2 分类标准

模型通过全连接层和 softmax 分类器来实现最后的分类输出,根据民航局新规中机场对延误的定义,航班实际离港时间与计划起飞时间之间的差值在 15 分钟以内,航班属于未延误,参考专家意见对于超出 15 分钟的再划分为不同的延误等级,因此,将延误时长 (T) 划分为五个等级,分别是 $T \leq 15$ 、 $15 < T \leq 60$ 、 $60 < T \leq 120$ 、 $120 < T \leq 240$ 和 $T > 240$,分类器根据划分方式进行延误等级预测。

4.3 实验结果分析

实验首先比较的是 CNN、LSTM、CNN-LSTM 以及文中使用的融合模型 CNN-LSTM-ATT 在数据集上的分类效果。CNN-LSTM-ATT 先使用 CNN 对特征之间的联系进行挖掘,再将寻找到关联的特征扩展成以为向量输入到 LSTM 中进行分类预测的操作,注意力机制对重要的信息给予更多的关注供模型学习。

CNN 模型设置 3 次卷积操作,卷积核长度设置为 2,两层最大池化层,参数设置为 2, LSTM 模型中 LSTM 层数为 2,第一层包含 128 个隐层通道,第二层包含 64 个隐层通道, CNN-LSTM 模型中层数以及参数设置按照文中提出的融合模型设置相同,实验中模型训练轮数设置为 85 次,每一次迭代批处理的样本量为 100。实验中观察到,融合模型相较独立模型耗时稍长,一次训练差值在 7 s 左右,训练耗时相差不大。

图 4 展示了四个模型的损失值对比和准确率对比,

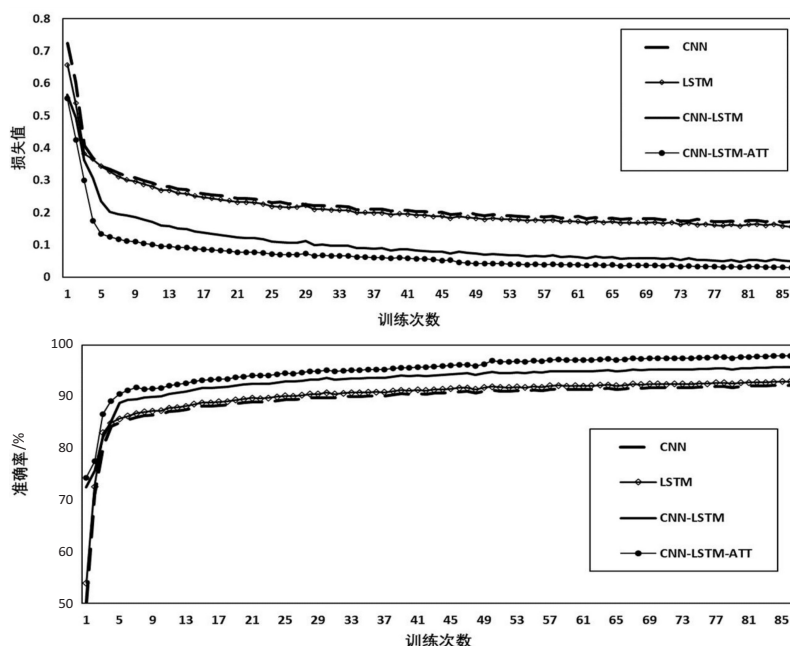


图 4 模型损失值和准确率对比

融合模型损失值收敛性能较好,损失值经过训练降到 0.1 左右。融合模型准确率比单一模型准确率高,单一模型准确率为 92% 左右,各种融合模型的准确率均能达到 96% 以上,其中提出的融合模型 CNN-LSTM-ATT 准确率高达 98.1%,相比单一模型提升了 6 个百分点。CNN 模型可以有效获取一条数据的关键特征,但难以发现时间序列之间的时序特征,场面上航班延误具有聚集性,机场运营数据按起飞时间排序后输入模型,CNN 模型难以捕捉每条数据的前后信息。LSTM 模型难以挖掘一条数据的内部相关性,数据内部的多项特征之间存在一定的联系,例如,天气现象和前一小时延误架次存在相关性。CNN-LSTM 模型经过训练准确率可以达到 96.2%,相比独立模型,分类效果得到了提升,但对于部分延误数据模型未能察觉其特殊性,模型没有给予延误数据特定的权重值进行标记,使得模型对这些数据难以做出准确的判断。

为验证提出的 PN 模型获取的动态特征对模型分类预测准确率的提升,设置对比实验,评估该特征在模型训练中的作用,分别将加入动态特征的和未加入动态特征的特征数据输入到 CNN-LSTM-ATT 预测模型。图 5 所示为两种输入特征下的模型准确率,两组实验都进行了 85 次训练,未添加动态特征的数据集在该融合模中最终获得了 93.9% 的准确率,实验中该数据集在进行了 50 次训练时,准确率趋于平稳,随后准确率便没有明显提升,表明模型难以对存在延误的特殊信息进行进一步的学习。图 6 显示了两种输入数据在模型训练后输出的混淆矩阵,后三种是存在明显延误情况的航班,模型对两种输入特征中不存在明显延误现象的航班能够准确分类,动态特征的输入给延误航班提供了更大的学习特性,让模型能够更加准确地识别信息。

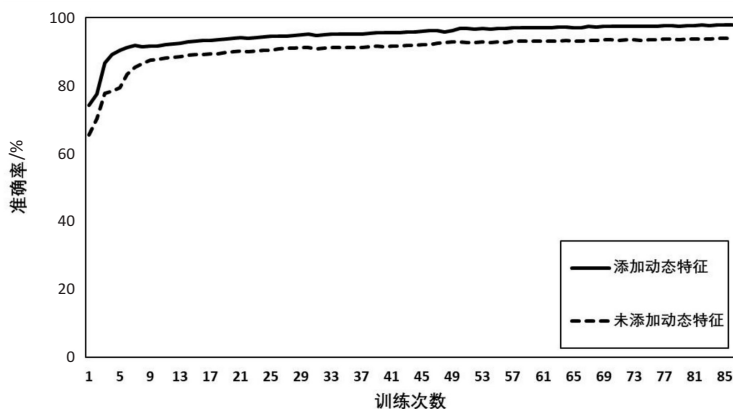


图5 不同输入特征的模型准确率

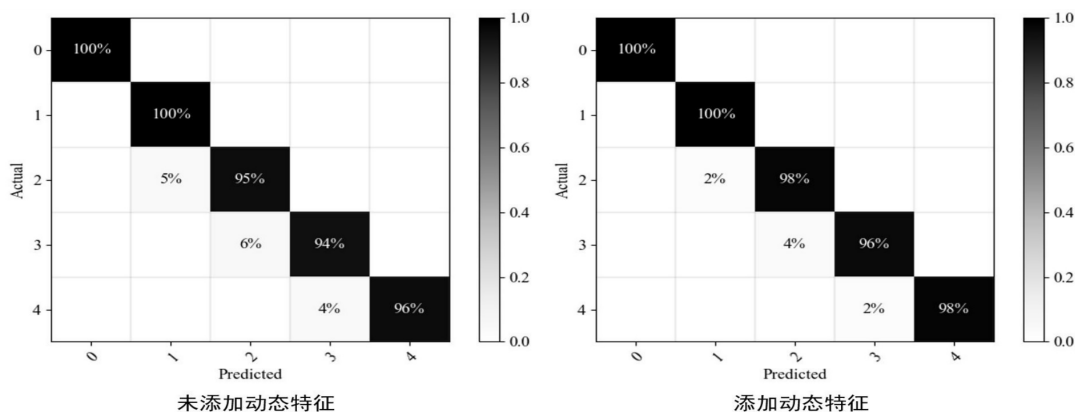


图6 不同输入特征的模型混淆矩阵

目前,对于保障作业时长优化,国内大型机场多采用人工统计和二八法对历史数据进行分析,从而调整作业时长来降低延误风险,但上述方法都存在滞后性和准确率低的问题。该文提出的方法通过国内某大型机场的真实数据集验证了其可行性,可将 Petri Net 模型纳入机场协同决策(Airport Collaborative Decision Making,ACDM)系统中,通过分析地面保障流程,依据

历史数据自我学习,动态更新预测模型的输入特征,为航空公司了解航班保障情况、采取流程优化措施提供有效参考。

5 结束语

该文提出了一种基于融合神经网络的航班延误情况分类预测模型,根据机场真实航班保障流程建立时

序 Petri Net 模型,获取保障流程关键作业,成为预测模型的微观动态特征。结合 CNN、LSTM 和注意力机制的优势,构建融合模型 CNN-LSTM-ATT,该模型在实验数据集上获得了 98.1% 的准确率,比其他独立预测模型的准确率高,模型性能更好。航班保障流程作为一个动态流程,特征数据中加入关键作业时长特征项,对模型的分类准确率的提升是有很大作用的,让模型对重要延误信息有了更加准确的判断。在未来的工作中,可以继续考虑场面多航班资源竞争问题对保障时长的影响,模型方面可以进一步优化模型结构,缩短训练时长,提高模型分类准确率,对分类边缘数据做出更好的分类判断,并将模型应用到更多领域的流程分析工作中。

参考文献:

- [1] WANG Y, LI Y. Complexity analysis on the influence factors of the flight delay risk based on SNA[J]. Open Journal of Social Sciences, 2020, 8(5): 54-71.
- [2] KABONGO P C, RAMOS T, LEITE A F, et al. A multi-agent planning model for airport ground handling management [C]//IEEE international conference on intelligent transportation systems. Rio de Janeiro: IEEE, 2016: 2354-2359.
- [3] 杨 珏. 基于遗传算法的机场特种车辆调度应用研究[J]. 计算机技术与发展, 2019, 29(3): 164-168.
- [4] WU W, CAI K, YAN Y, et al. An improved SVM model for flight delay prediction [C]//2019 IEEE/AIAA 38th digital avionics systems conference (DASC). San Diego: IEEE, 2019: 1-6.
- [5] JIANG Y, MIAO J, ZHANG X, et al. A multi-index prediction method for flight delay based on long short-term memory network model [C]//2020 IEEE 2nd international conference on civil aviation safety and information technology (IC-CASIT). Weihai: IEEE, 2020: 159-163.
- [6] 陆继翔, 张琪培, 杨志宏, 等. 基于 CNN-LSTM 混合神经网络模型的短期负荷预测方法[J]. 电力系统自动化, 2019, 43(8): 131-137.
- [7] ALGULIYEV R M, ALIGULIYEV R M, ABDULLAYEVA F J. The improved LSTM and CNN models for DDoS attacks prediction in social media[J]. International Journal of Cyber Warfare and Terrorism, 2019, 9(1): 1-18.
- [8] CAO K, KIM H, HWANG C, et al. CNN-LSTM coupled model for prediction of waterworks operation data[J]. Journal of Information Processing Systems, 2018, 14(6): 1508-1520.
- [9] 孙本旺, 田 芳. 基于深度学习算法的藏文微博情感计算研究[J]. 计算机技术与发展, 2019, 29(10): 55-58.
- [10] LIVIERIS I E, PINTELAS E, PINTELAS P. A CNN-LSTM model for gold price time-series forecasting [J]. Neural Computing and Applications, 2020, 32: 17351-17360.
- [11] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Advances in neural information processing systems. Long Beach: [s. n.], 2017: 5998-6008.
- [12] ZHU X. A Petri-net modeling approach for airport apron traffic dynamics [C]//2017 36th Chinese control conference. Dalian: TCCT, 2017: 2332-2337.
- [13] DAVIDRAJUH R, RONG C. Finding the pivotal elements for modularization of petri nets [C]//2019 international conference on advanced mechatronic systems (ICAMEchS). Shiga: IEEE, 2019: 92-97.
- [14] LUO Q, CHEN Y, CHEN L, et al. Research on situation awareness of airport operation based on petri nets [J]. IEEE Access, 2019, 7: 25438-25451.
- [15] 汤新民, 朱新平. Petri 网原理及其在民航交通运输工程中的应用[M]. 北京: 中国民航出版社, 2014.
- [16] GU J, WANG Z, KUEN J, et al. Recent advances in convolutional neural networks [J]. Pattern Recognition, 2018, 77: 354-377.
- [17] PASCANU R, MIKOLOV T, BENGIO Y. On the difficulty of training recurrent neural networks [C]//Proceedings of the 30th international conference on machine learning - volume 28. Atlanta: JMLR. org, 2013: III-1310 - III-1318.