

基于图卷积网络的热轧带钢轧制力预测

李维刚^{1,2}, 刘玮汲², 谢璐², 赵云涛^{1,2}

(1. 武汉科技大学冶金自动化与检测技术教育部工程研究中心, 湖北 武汉 430081;

2. 武汉科技大学信息科学与工程学院, 湖北 武汉 430081)

摘 要: 热连轧生产为多钢种、多规格混杂的带钢连续轧制过程, 现有的机器学习方法不能考虑各带钢层的影响, 将各带钢的轧制力预测过程视为独立的而不是关联的, 这种做法不符合实际情况。提出一种预测带钢轧制力的梯度提升树-图卷积神经网络 (gradient boosting decision tree-graph convolutional networks, GBDT-GCN) 模型。首先, 构建用于轧制力预测的带钢关系图结构, 将数据集中的每块带钢作为图结构中的节点, 根据带钢的轧制时序、层别关系生成各带钢节点之间的连接边, 将连续轧制、相同层别的带钢关联起来; 接着, 将图结构输入结构调整后的 GCN 模型, 采用平均绝对误差作为损失函数进行模型训练, 采用 GBDT 对轧制力的影响因素进行重要性排序, 并根据 GCN 模型的预测精度变化筛选出重要的因子作为最终的节点特征向量。最后, 利用国内某热连轧机组的实际生产数据进行试验验证, 结果表明, GBDT-GCN 模型在测试集上的平均绝对误差为 405.6 kN, 相对误差在 $\pm 10\%$ 以内的数据所占比例为 91.5%, 相较于传统 SIMS 模型、RF 随机森林算法、MLP 多层感知机模型, 利用带钢关系图结构预测轧制力的 GBDT-GCN 模型具有更高的预报精度。

关键词: 热连轧带钢; 轧制力预测; 特征选择; 梯度提升树; 图卷积神经网络

文献标志码: A **文章编号:** 0449-749X(2023)03-0089-08

Rolling force prediction of hot rolled strip by graph convolutional networks

LI Wei-gang^{1,2}, LIU Wei-ji², XIE Lu², ZHAO Yun-tao^{1,2}

(1. Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, Hubei, China; 2. School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, Hubei, China)

Abstract: Hot strip rolling is a continuous rolling process of strip steel with multiple steel grades and specifications, however the current machine learning methods could not consider the influence of strip layers, and treat the rolling force prediction process as an independent process of each strip rather than a related process, which is not in line with the actual situation. A gradient boosting decision tree-graph convolutional networks (GBDT-GCN) model is proposed to predict the rolling force of strip steel. Firstly, a strip relational graph structure for rolling force prediction was constructed. Each strip in the data set was regarded as a node in the graph structure, and the connecting edges were generated according to the rolling sequence and layer relationship of the strip. The same layer strips or the strips which were continuously rolled are associated by graph edges. Then, the graph structure was input into the adjusted GCN, and the mean absolute error was used as the loss function for model training, and the GBDT was utilized to rank the influencing factors of rolling force, the important factors were selected as the final node feature vector according to the change of prediction accuracy of GCN model. Finally, the actual production data of a domestic hot strip mill are used for experimental verification. The results show that the average absolute error of the GBDT-GCN model on the test set is 405.6 kN, and the proportion of data with relative error within $\pm 10\%$ is 91.5%. Compared with the traditional SIMS model, RF random forest algorithm and MLP multi-layer perceptron model, the GBDT-GCN model of predicting rolling force by strip graph structure has higher prediction accuracy.

Key words: hot strip rolling; rolling force prediction; feature selection; gradient boosting decision tree; graph convolutional networks

轧制力预测是热连轧过程控制的关键一环,轧制力预测不准确将严重影响带钢的产品质量与轧制稳定性^[1]。传统的轧制力模型大多建立在轧制理论基础之上,一些常用假设与实际情况存在差异,如轧辊压扁形状假设、平面变形假设等,这些问题限制了模型的预测精度^[2]。另外,热连轧过程数据波动和检测误差大、信噪比低、分布不均匀,且轧制力与成分、工艺各参数的作用机理复杂;不同钢种、规格的带钢轧制特性差异明显,实际生产过程经常发生较为复杂的层别切换的交叉轧制。单靠机理模型进行轧制力预测,会忽略很多现场工艺和设备状态的变化。为此,实际在线使用的轧制力模型都会采用基于层别数据的模型自学习功能^[3],并引入各种自学习系数以适应复杂多变的轧制工况^[4],但仍然存在实时性较差和准确性较低的问题^[5]。

随着计算机技术的发展,利用机器学习、深度学习算法提升轧制力预测精度成为研究的热点^[6]。陈治明等^[7]在最小二乘支持向量机模型的基础上,提出了一种结合遗传算法的变尺度混沌优化方法,预测结果的平均误差率显著降低。冀秀梅等^[8]提出了基于极限学习机(ELM)的综合神经网络轧制力预测模型,通过化学成分计算出基准变形抗力后,将结果作为神经网络的输入变量进行轧制力预测,该方法使某钢种头部厚度命中率提高了 1.9%。刘明华等^[9]基于板带轧制数据和支持向量回归(SVR)模型对轧制力进行预测,采用主成分分析法(PCA)降低输入变量的维数,同时使用改进粒子群优化(IPSO)算法对 SVR 模型参数进行优化。马威等^[10]构建了一种深度前馈神经网络轧制力模型,运用融入 Mini-Batch 策略的 Adam 优化算法进行参数寻优,采用参数惩罚和 Dropout 正则化策略提高了模型的泛化能力。然而,上述研究仅仅考虑了单块带钢的轧制力预报,忽略了各带钢之间的轧制时序、层别等关联关系,严重限制了模型的适用范围。

针对上述问题,本文提出了一种预测热连轧带钢轧制力的 GBDT-GCN (gradient boosting decision tree-graph convolutional networks) 模型。首先,根据带钢的轧制时序、层别关系构建用于轧制力预测的图结构,利用关系图将连续轧制、相同层别的带钢关联起来;然后,采用结构调整后的 GCN 回归模型进行轧制力预测,并结合 GBDT 重要性排序和轧制力模型预测结果进行特征选择,剔除部分冗余特征后作为模型最终的输入参数;最后,通过钢厂的实际数据验证了 GBDT-GCN 模型的可行性和准确性。

1 图卷积神经网络

GCN 是一种利用卷积来提取空间特征的图神经网络^[11]。传统的卷积神经网络(convolutional neural network, CNN)只能对有固定结构的数据进行建模和预测,而真实世界中很多数据例如社交网络、知识图谱、推荐系统都是不规则的图谱结构, CNN 无法处理这些非欧氏空间结构。因此, GCN 被提出并用于处理非结构化的数据, GCN 基于信息传播机制引入卷积,每一个节点利用相互交换信息来更新自己的节点状态,直到达到某一个稳定值。通过提取空间特征来学习节点表示的方法,使得 GCN 在节点分类、图分类、交通流量预测等领域中都取得较好效果。

图结构定义为 $G = (U, E)$, U 为节点的集合, E 为边的集合,边可描述节点之间的关系。将特征矩阵 $\mathbf{X}^l \in R^{n \times d}$ 和邻接矩阵 $\mathbf{A} \in R^{n \times n}$ 作为 GCN 的输入, n 为节点数量, d 为节点特征维度, l 为网络的第 l 层。输出特征矩阵为 $\mathbf{X}^{l+1} \in R^{n \times d'}$, d' 为新的特征维度,一个 GCN 卷积层可以表示为

$$\mathbf{X}^{l+1} = h(\hat{\mathbf{A}}\mathbf{X}^l\mathbf{Z}^l) \quad (1)$$

式中: $\hat{\mathbf{A}}$ 为归一化后的邻接矩阵 $\mathbf{A}$; $\mathbf{Z} \in R^{d \times d'}$ 为待学习的变换矩阵; $h(\cdot)$ 为激活函数 LeakyReLU。

GCN 的学习方式遵循逐层学习,所有节点同步更新, GCN 的本质是在节点间传播信息来更新节点表示。因此, GCN 可在图结构引导下挖掘节点特征和标签之间的共生关系^[12]。分类问题和回归问题本质上是相同的,都是建立一个函数映射关系 $g(\cdot)$, CNN 在处理回归问题时,通过改变基础网络结构,并将损失函数替换成回归问题的损失函数, CNN 便有很好的回归预测效果。从另一个角度来看, GCN 的变换矩阵 $\mathbf{Z}$ 可以看成 n 个一维卷积核,卷积将节点特征维度由 d 维缩放到 d' 维,说明 GCN 和 CNN 在结构上具有一定的相似性,类比 CNN 处理回归问题的方案,利用 GCN 进行回归预测也有较好的可行性^[13]。

2 GBDT-GCN 轧制力预测模型

要建立轧制力预测的 GCN 模型,需要先构建相应的图结构。本文将用于建模的各块带钢作为图结构中的节点,将轧制力的影响因素作为节点特征向量,其维度由 GBDT 特征选择决定,所有带钢的特征向量构成节点特征矩阵;用图结构的边描述带钢之间的联系,根据带钢的轧制时序、层别关系构成

邻接矩阵。特征矩阵和邻接矩阵得到后,可建立轧制力预测的 GCN 回归模型。

2.1 试验数据的获取

从国内某 1 780 热连轧机组的 L2 在线数据库中获取试验数据,共收集 12 966 块带钢数据,涵盖 65 个钢种大类、815 个细分层别。本文以 F6 机架为例进行建模分析。目前,公认最适用于热连轧带钢轧制力预报的数学模型是基于 OROWAN 变形

压力平衡理论的 SIMS 模型,根据大部分实际现场使用的 SIMS 模型,轧制力模型的初始输入特征组成见表 1。选择入口厚度、出口厚度、轧制温度、带钢宽度、带钢速度、机架前张力、机架后张力、轧辊直径,以及各种主要的化学成分作为模型的初始输入特征,一共包含 31 个变量,输出为机架的实测轧制力。由于人工选择的初始特征较多,存在部分无效特征,所以后面要进行特征筛选。

表 1 轧制力模型的初始输入特征
Table 1 Initial input features of rolling force model

序号	输入特征	特征名	最小值	最大值
1	入口厚度/mm	ENTRY_THICK	1.70	15.36
2	出口厚度/mm	EXIT_THICK	1.38	13.64
3	轧制温度/℃	ENTRY_TEMP	843.52	992.06
4	带钢宽度/mm	STRIP_WIDTH	845.41	1 663.95
5	带钢速度/(m·s ⁻¹)	STRIP_SPEED	2.49	12.73
6	入口张力/MPa	ENTRY_TENSION	7.20	46.80
7	出口张力/MPa	EXIT_TENSION	8.40	62.40
8	轧辊直径/mm	ROLL_DIAM	642.59	687.76
9	w(C)/%	CARBON	0.01	0.23
10	w(Si)/%	SILICON	0	2.18
11	w(Mn)/%	MANGANESE	0	2.80
12	w(P)/%	PHOSPHORUS	0	0.13
13	w(S)/%	SULPHUR	0	0.03
14	w(Cu)/%	COPPER	0	0.30
15	w(Ni)/%	NICKEL	0	0.15
16	w(Cr)/%	CHROMIUM	0	0.59
17	w(As)/%	ARSENIC	0	0.01
18	w(Sn)/%	TIN	0	0.12
19	w(Nb)/%	NIOBIUM	0	0.05
20	w(V)/%	VANADIUM	0	0.12
21	w(Al)/%	ALUMINIUM	0	0.77
22	w(Ti)/%	TITANIUM	0	0.11
23	w(Mo)/%	MOLYBDENUM	0	0.23
24	w(B)/%	BORON	0	0.01
25	w(W)/%	TUNGSTEN	0	0.01
26	w(Ca)/%	CALCIUM	0	0.01
27	w(H)/%	HYDROGEN	0	0.01
28	w(O)/%	OXYGEN	0	0.01
29	w(N)/%	NITROGEN	0	0.01
30	w(SOL Al)(酸溶铝)/%	SOL_ALUMINIUM	0	0.76
31	w(Sb)/%	ANTIMONY	0	0.06

2.2 模型输入特征的选择

由于初始挑选出来的轧制力模型输入特征较多, 本文采用 GBDT 算法对各种输入特征进行重要性排序。GBDT 算法是以 CART 回归树为基模型的 boosting 算法, GBDT 输出的结果是由若干棵决策树模型累加而成, 每一棵决策树都是对之前决策树组合预测残差的拟合, 累加输出和残差拟合增加了 GBDT 的泛化能力和预测精度, 特征的重要性评分可直接从拟合模型中得到, GBDT 算法流程如图 1 所示。

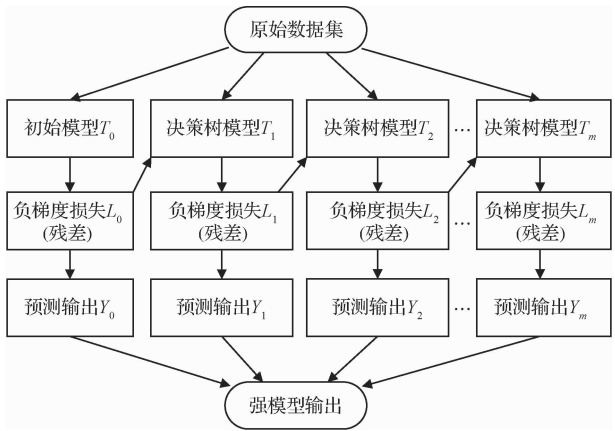


图 1 GBDT 算法流程

Fig. 1 Simple flow chart of GBDT algorithm

原始数据拟合所有 GBDT 的决策树后, 通过计算特征 q 在单棵树中重要性的平均值, 来衡量特征 q 在 GBDT 的全局重要性^[14]。特征 q 在单棵树 T 中的重要性评分 Q_q 为

$$Q_q^2(T) = \sum_{t=1}^{J-1} u_t^2 1 \quad (v_t = q) \tag{2}$$

式中: $J-1$ 为树的非叶子节点数量 (构建的树都是具有左右孩子的二叉树); v_t 为和节点 t 相关联的特征; u_t^2 为节点 t 分裂之后平方损失的减少值。

在 GBDT 中树的数量为 M , 则特征 q 的全局重要性 ϵ_q 可表示为

$$\epsilon_q = \frac{1}{M} \sum_{m=1}^M Q_q^2(T_m) \tag{3}$$

式中: T_m 为第 m 棵树, $m=1, 2, \dots, M$ 。

对所有特征计算全局重要性评分后进行排序, 利用 GCN 回归模型剔除重要性评分排在最后的部分特征, 保留下来的特征就可作为最终的带钢节点特征向量。

2.3 轧制力模型的构建

2.3.1 图结构

轧制过程数学模型大多建立在轧制理论的基础

之上, 由于计算速度以及应用性能的限制, 目前在轧线上使用的模型均为理论模型的简化版, 这限制了轧制模型的设定精度, 基于层别数据的自学习为提高轧制模型的计算精度提供了切入点^[2]。传统轧制模型从历史数据中学习得到模型的自学习系数, 并将其应用在后续带钢的设定计算中, 这种方法能减弱因本体模型设定不准确带来的误差。因此, 每块带钢的轧制参数设定不是独立的而是关联的, 这种关联主要体现在前后轧制带钢的联系和同一层别的联系。

前后轧制带钢的联系主要表现为连续轧制的带钢具有相似的设备工艺条件与轧制状态, 如工作辊的热膨胀、磨损、加热炉温度等, 这些因素使得连续轧制的带钢之间存在一定的相似性与时序关系。

同一层别带钢的联系主要表现为相同钢种、规格的带钢在加热制度、物性参数、化学成分、轧制工艺、带钢性能等方面具有共性, 相同层别的带钢具有相似的轧制特性。

根据上述分析, 将数据集中的每块带钢作为图结构的节点, 根据带钢的轧制时序与层别关系获得图结构的连接边。本文提出图结构的更新规则: (1) 时序更新规则, 带钢的连续轧制使温度设定与辊缝设定在时序上相关, 前一块带钢与后一块带钢存在关联, 将前者到后者之间的连接作为图结构的边; (2) 层别更新规则, 相同层别的带钢具有相似的轧制特性, 越近邻的带钢联系越密切, 追加同层别轧制的前 2 块带钢到本块带钢的连接作为图结构的 2 条边, 较近邻的带钢给予较大的权重系数。

根据上述 2 条规则, 构建描述 n 个带钢节点之间连接关系的邻接矩阵 $\mathbf{A}^{n \times n}$, 步骤如下。

Step1: 连续轧制的带钢样本集按时间先后排列为 $\mu_1, \mu_2, \dots, \mu_n$;

Step2: 初始化带钢邻接矩阵 $\mathbf{A}^{n \times n}$ 的元素全为 0, a_{ij} 为 $\mathbf{A}^{n \times n}$ 矩阵内第 i 行第 j 列对应的元素;

Step3: 设置循环 $i = 1, 2, \dots, n-1$, 循环执行 $a_{i(i+1)} = 1$, 此时轧制时序更新规则结束;

Step4: 按轧制时间顺序从后往前建立带钢样本之间的层别联系, 初始化外循环 $j = n$;

Step5: 设置内循环 $i = j-1, j-2, \dots, 1$, 当循环到 μ_i 的层别和 μ_j 的层别相同时, 执行 $a_{ij} = a_{ij} + 1$, 继续执行内循环直到 μ_i 与 μ_j 的层别再次相同时, 执行 $a_{ij} = a_{ij} + \frac{1}{2}$, 并跳出内循环;

Step6: 带钢 μ_j 的边构建完成后, 接着构建前一块带钢 μ_{j-1} 的边, 即执行外循环 $j = j-1$, 并跳到

Step5,直到 $j = 1$ 时退出整个循环。

带钢关系图结构的示意如图 2 所示,图中每个节点代表 1 块带钢,相同颜色的节点表示相同层别的带钢。可见,相邻轧制的带钢与相同层别的带钢之间存在连接边。

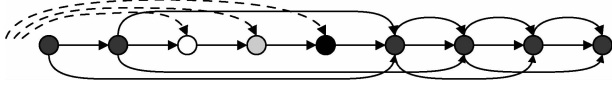


图 2 带钢关系图结构的示意

Fig. 2 Graph diagram for strip steel relationship

2.3.2 GCN 回归模型

GCN 具有在节点之间传播信息的特性,在不断训练更新模型的权重后,能够输出连续的节点值,所以 GCN 具有解决回归问题的能力。热连轧轧制力预测对输入特征十分敏感,选择 GCN 网络层数时必须考虑过平滑缺陷。当 GCN 网络过深时,节点之间的信息传递会使得各个节点的差异性减弱,节点特征趋于一致对于参数敏感的轧制力预测是不利的,输出特征过平滑、节点差异性太小会导致模型输出范围变小、预测能力变弱,因此选择只含有 1 个隐藏层、1 个回归层的 2 层 GCN 模型作为轧制力预测模型。

GCN 的输出节点是自身节点特征和相邻节点特征的聚合,聚合比例由邻接矩阵决定。在分类任务中,为了让同类别节点的输出趋于一致,直接将邻接矩阵代入 GCN,聚合时相邻节点占较高比例,但这样会导致网络输出的过平滑,无法保证节点预测值的连续性,不适用于轧制力预测这类回归问题。文献[15]提出加权减弱过平滑的方法,将邻接矩阵乘以一个权重系数 $\theta \in (0,1)$,以降低聚合时相邻节点的比例,使信息传递更加关注节点自身。受此启发,本文提出一种适用于回归问题的邻接矩阵加权方法,即

$$\tilde{\mathbf{A}} = \frac{\mathbf{A}}{d} + \mathbf{I}_n \quad (4)$$

式中: $\tilde{\mathbf{A}}$ 为加权后的邻接矩阵; $\mathbf{A}$ 为原始的邻接矩阵; d 为输入节点的特征维度; $\mathbf{I}_n$ 为单位矩阵。

因为带钢之间的轧制时序、层别关系是单向的,相应的图结构是有向图,因此 $\mathbf{A}$ 、 $\tilde{\mathbf{A}}$ 都是非对称矩阵。

可见, $\tilde{\mathbf{A}}$ 与特征维度 d 相关,当节点特征维度较大时,自身节点特征对输出的影响较大,而相邻节点对输出的影响较小,这符合轧制力预测时主要依赖于自身节点的特征,而其他带钢的影响相对较小。

综上所述,当输入特征矩阵 $\mathbf{X}$ 为 d 维、隐藏层为 d' 维时,非对称矩阵 $\tilde{\mathbf{A}}$ 相应的 GCN 回归模型可表示为

$$\begin{aligned} f(\mathbf{X}, \tilde{\mathbf{A}}) &= \hat{\mathbf{A}} \cdot \text{LeakyReLU}(\hat{\mathbf{A}} \mathbf{X} \mathbf{Z}_0^{d \times d'}) \mathbf{Z}_1^{d' \times 1} \\ \hat{\mathbf{A}} &= (\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}})^T \\ \tilde{\mathbf{D}}_{ii} &= \sum_j \tilde{\mathbf{A}}_{ij} \end{aligned} \quad (5)$$

式中: $\mathbf{Z}_0$ 为第 1 个变换矩阵; $\mathbf{Z}_1$ 为第 2 个变换矩阵; $\hat{\mathbf{A}}$ 为归一化后的特征矩阵; $\tilde{\mathbf{D}}$ 为度矩阵; $\tilde{\mathbf{D}}_{ii}$ 为 $\tilde{\mathbf{D}}$ 的第 i 行第 i 列; $\tilde{\mathbf{A}}_{ij}$ 为 $\tilde{\mathbf{A}}$ 的第 i 行 j 列。

本文采用平均绝对误差作为 GBDT-GCN 模型的损失函数,训练过程动态调整学习率。

2.4 评价指标

本文选取平均绝对误差 ϵ (MAE)、均方根误差 σ (RMSE)、决定系数 R^2 和命中率 P_T 作为模型的评价指标。

$$\epsilon = \frac{1}{K} \sum_{k=1}^K |f_k - y_k| \quad (6)$$

$$\sigma = \sqrt{\frac{1}{K} \sum_{k=1}^K (f_k - y_k)^2} \quad (7)$$

$$R^2 = 1 - \frac{\sum_{k=1}^K (f_k - y_k)^2}{\sum_{k=1}^K (y_{\text{mean}} - y_k)^2} \quad (8)$$

$$P_T = \frac{\sum_{k=1}^K \left(\frac{|f_k - y_k|}{y_k} \leq 10\% \right)}{K} \times 100\% \quad (9)$$

式中: f_k 为轧制力模型的预测值; y_k 为轧制力的实测值; y_{mean} 为 K 块带钢实测轧制力的均值; K 为带钢块数。

3 试验结果与分析

试验选取中国宝武钢铁集团某 1 780 热连轧机组的实际生产数据,收集 3 个月共 12 966 块带钢的数据。将数据按照 7 : 3 的比例划分训练集和测试集,人工选择的初始节点特征数为 31 个,对原始特征进行均值方差归一化。将数据集集中的带钢按轧制时序、层别关系构建轧制力预测用的图结构,并输入 GCN 模型,使用 Adam 优化器进行参数寻优,设置初始的学习率为 0.1。

3.1 GCN 网络结构选择

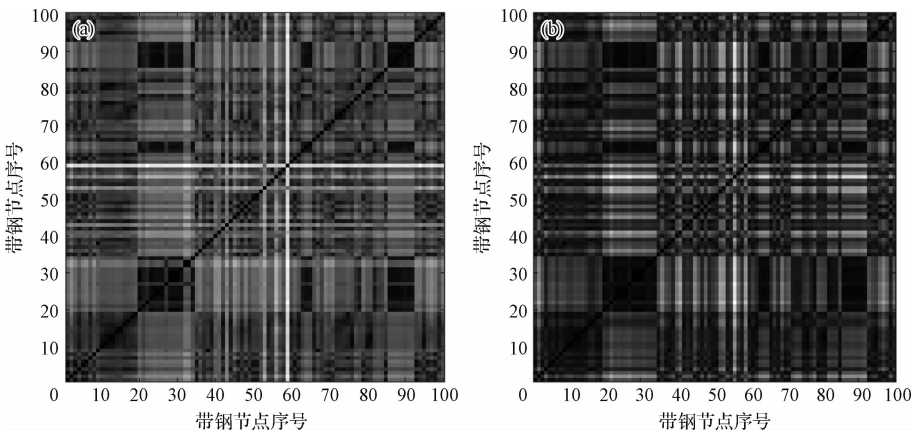
如果使用过多层数的 GCN 深度堆叠,节点的区别性会变得越来越弱,预测的效果会急剧下降,本文选择 2 层结构(1 个隐藏层、1 个回归层)的 GCN 网络,设置隐藏层节点维度为 16。为了验证 2 层

GCN 的效果,引入 3 层 GCN 作为对比,设置其第 1、第 2 个隐藏层的节点维度分别为 24、16。当训练集 ϵ 达到 800 时网络训练停止,分别拉取 2 个模型前 100 个的输入在最后 1 个隐藏层的节点特征向量。定义节点的欧式距离 ϕ 为差异度,有

$$\phi(\boldsymbol{\alpha},\boldsymbol{\beta})=\sqrt{\sum_{\lambda=1}^{16}(\alpha_{\lambda}-\beta_{\lambda})^2}\tag{10}$$

式中： $\boldsymbol{\alpha}$ 、 $\boldsymbol{\beta}$ 为节点特征向量； λ 为该特征在特征向量的第 λ 维； α_{λ} 、 β_{λ} 为第 λ 维的节点特征值。

根据式(10)分别求得 2 个 100×100 的差异度



(a) 2 层 GCN 节点差异度； (b) 3 层 GCN 节点差异度

图 3 GCN 节点差异度对比

Fig. 3 Comparison diagram of GCN node difference

将分类网络改成回归网络时,文献[16]将主干网络作为编码器,最后 1 层加入解码器输出连续值。为了验证 GCN 本身就有很好的回归能力,将 2 层 GCN 后加入全连接层(fully connected layer,FC)得到“GCN+FC”模型,“GCN+FC”回归模型和“纯 GCN”回归模型在测试集上的精度对比见表 2。

表 2 网络结构对比试验

Table 2 Comparative experiment of network structure			
网络结构	维度变换	ϵ/kN	σ/kN
GCN+FC	GCN:31→16→10,FC:10→1	509.2	697.2
GCN(regression)	GCN:31→16→1	428.5	583.5

由表 2 可知,直接将 GCN 的最后 1 层作为回归网络的输出层,2 个精度指标都比“GCN+FC”要好,所以本文选取 2 层 GCN 回归模型作为预测模型。另外,由于上述对比试验的输入节点都为未进行特征优化的初始节点,可将上述结果作为 GBDT 特征选择的对比基础。

矩阵,归一化后绘制热力图,颜色越深代表节点之间的差异度越小,节点过平滑越明显,如图 3 所示。2 层 GCN 的输出节点差异度明显大于 3 层 GCN,3 层 GCN 节点过平滑效应明显。虽然 GCN 能通过聚合相邻节点进行预测,但 3 层 GCN 的网络层数过高,导致节点差异性显著减弱,如果继续堆叠网络,节点的表示向量将趋于一致。因此,GCN 用于轧制力预测时,2 层 GCN 就能取得较好的预测效果,过多的网络层数不仅会增加模型的复杂度,还会导致预测精度降低。

3.2 特征重要性排序

GBDT 算法设置 CART 回归树为 150 棵,学习率为 0.1,树深度取默认值 3。由于轧制力数据分布极不均衡,因此对 GBDT 进行初始化时,使用能包含更多样本分布信息的中位数而不是平均值,该措施能减小极值样本对模型的影响。将人工初步选取的节点特征代入 GBDT 算法中,输出特征重要性的排序见表 3。

表 3 GBDT 轧制力相关特征重要性排序

Table 3 Order of importance of GBDT rolling force related features		
排序	输入特征	重要性评分
1	带钢宽度	0.117 2
2	入口厚度	0.097 2
3	出口厚度	0.091 5
4	带钢速度	0.070 5
5	轧制温度	0.065 9
6	轧辊直径	0.062 6
7	$w(\text{C})$	0.040 6
8	$w(\text{Mn})$	0.038 7

续表		
排序	输入特征	重要性评分
9	$w(\text{Si})$	0.033 7
10	$w(\text{P})$	0.033 3
11	$w(\text{Ti})$	0.029 0
12	$w(\text{SOL Al})$	0.028 5
13	出口张力	0.028 1
14	$w(\text{Nb})$	0.026 2
15	入口张力	0.025 1
16	$w(\text{Al})$	0.021 0
17	$w(\text{S})$	0.021 0
18	$w(\text{Cu})$	0.019 0
19	$w(\text{N})$	0.018 6
20	$w(\text{B})$	0.017 1
21	$w(\text{Cr})$	0.015 7
22	$w(\text{Ni})$	0.013 5
23	$w(\text{O})$	0.012 9
24	$w(\text{H})$	0.012 2
25	$w(\text{As})$	0.012 0
26	$w(\text{Mo})$	0.011 6
27	$w(\text{Sb})$	0.011 4
28	$w(\text{V})$	0.007 8
29	$w(\text{W})$	0.007 5
30	$w(\text{Ca})$	0.006 6
31	$w(\text{S})\text{n}$	0.002 4

3.3 模型预测结果分析

根据表 3 的 GBDT 重要性排序结果,按重要性评分由小到大依次删除特征,保留的特征作为输入节点特征代入 GCN 回归模型中,依次得到测试集精度。为了方便进行特征选择,定义评价标准 $A_{cc}=0.6\epsilon+0.4\sigma$,结果见表 4。

表 4 GCN 精度随删除特征的变化

Table 4 Variation of GCN accuracy with deleted features			
删除特征个数	ϵ/kN	σ/kN	A_{cc}/kN
0	428.5	583.5	490.5
1	427.9	586.4	491.3
2	423.4	586.7	488.7
3	411.0	565.7	472.9
4	414.8	572.4	477.9
5	416.7	563.0	475.2
6	405.1	560.2	467.1
7	414.6	553.4	470.1
8	402.5	558.8	465.0
9	405.6	552.0	464.1
10	424.3	579.8	486.5
11	443.1	603.1	507.1

由表 4 可知:(1)当删除 8 个特征时 $\epsilon=402.5\text{ kN}$ 最小,删除 9 个特征时 $\sigma=552.0\text{ kN}$ 最小,删除 9 个特征时 $A_{cc}=464.1\text{ kN}$ 最小;(2)当删除的特征

超过 9 个以后,模型的精度有所降低,这说明如果删除的变量过多,就会丢失一些有效数据,导致模型的预测性能变差。

因此,选择删除 9 个特征,保留的 22 个特征作为 GCN 模型的输入,优化后的输入特征对应表 3 中排在前面的 22 个因素。此时,对比未进行特征筛选的 GCN 模型,GBDT-GCN 模型的 ϵ 降低了 22.9 kN, σ 降低了 31.5 kN。

GBDT-GCN 模型在测试集上的预测轧制力和实测轧制力的对比如图 4 所示,实线表示模型预测值和实测值完全一致,虚线为相对误差等于 $\pm 10\%$ 的边界线,超出边界的点数量较少且偏差不大。可见,在不划分钢种、规格的前提下,GBDT-GCN 对于轧制力跨度大的数据集也有较好的预测效果。

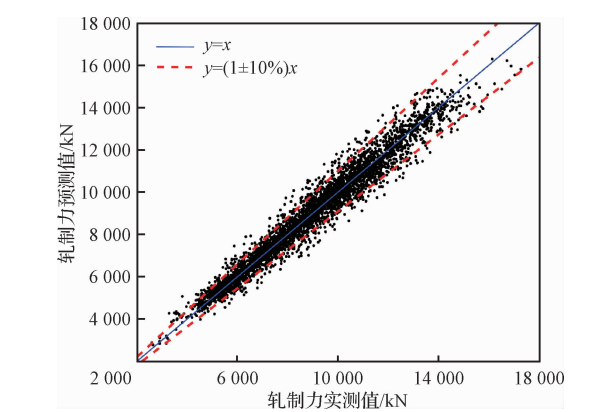


图 4 轧制力预测值和实测值的散点图
Fig. 4 Scatter plot of predicted and actual rolling force

为了验证 GBDT-GCN 模型的准确性,引入其他主流的轧制力模型进行对比。试验选取 SIMS 模型、MLP 多层感知机模型^[10]、RF 随机森林模型,与 GBDT-GCN 模型进行对比,通过计算各个模型在测试集上的命中率 P_T 、 ϵ 、 σ 和 R^2 这 4 个指标来评价模型的预测效果,对比结果见表 5。其中,SIMS 模型直接采用某 1 780 机组在线模型的预测结果数据,该模型包括了在线自学习功能。

表 5 不同模型的精度对比

Table 5 Precision comparison of different model				
模型	$P_T/\%$	ϵ/kN	σ/kN	R^2
SIMS	84.2	475.7	629.2	0.938
MLP	83.5	489.5	648.2	0.937
RF	86.1	463.1	660.8	0.934
GCN	89.6	428.5	583.5	0.957
GBDT-GCN	91.5	405.6	552.0	0.963

由表 5 可知:(1)对比未经 GBDT 优化特征的 GCN 模型,GBDT-GCN 模型精度有一定提升,这是因为剔除了冗余特征后,模型保留了更具代表性的特征变量;(2)GBDT-GCN 相对于 SIMS、MLP、RF 等模型的精度有较显著的提升,这是因为对于多钢种、多规格的实际轧制过程,GBDT-GCN 模型考虑了相邻带钢、同层别带钢的影响,利用带钢关系图实现了轧制力的联合预测,极大提升了模型的预测精度。这也是 GBDT-GCN 模型与其他机器学习模型的区别,新模型融合了更多的现场工艺规则,可实现对热连轧多钢种、多规格轧制模式下轧制力的精准预测,有广泛的理论研究与应用价值,尤其在热连轧带钢质量诊断分析领域应用前景广阔。

4 结 论

(1)构建用于轧制力预测的图结构。将数据集中的带钢作为图结构的节点,将轧制力的影响因素作为节点特征向量,所有带钢的特征向量构成节点特征矩阵;根据带钢的轧制时序、层别关系生成图结构的连接边,将连续轧制、相同层别的带钢节点关联起来,获得反映带钢节点关系的邻接矩阵。

(2)建立轧制力预测的 GCN 模型。将上述图结构输入结构调整后的 GCN 模型,采用平均绝对误差作为损失函数进行模型训练,利用加权方法使邻接矩阵适合回归问题;采用 GBDT 梯度提升树算法对初始的模型输入特征进行重要性排序,再结合 GCN 模型的预测精度进行节点特征筛选。

(3)通过实际钢厂的数据验证表明,GBDT-GCN 模型在相对误差区间±10%内轧制力预测占比为 91.5%,较传统的 SIMS 模型,以及 MLP、RF 等主流机器学习模型,GBDT-GCN 具有更高的预测精度;且新模型融合了更多的现场工艺规则,有广泛的理论研究与应用价值。

参考文献:

[1] 李海军,徐建忠,王国栋. 热轧带钢精轧过程高精度轧制力预测模型[J]. 东北大学学报(自然科学版),2009,30(5):669. (LI Hai-jun,XU Jian-zhong,WANG Guo-dong. High precision rolling force prediction model for hot rolled strip finishing [J]. Journal of Northeast University (Natural Science), 2009,30 (5):669.)

[2] 李维刚,邓肯,赵云涛,等. 基于连续曲面的轧制模型自学习方法[J]. 钢铁,2017,52(12):61. (LI Wei-gang,DENG Ken,ZHAO Yun-tao,et al. Self-learning method of rolling model based on continuous surface[J]. Iron and Steel, 2017, 52

(12):61.)

[3] 赵文姣,闫洪伟,杨枕,等. 提高轧制力设定精度的一种离线优化方法[J]. 中国冶金,2017,27(8):19. (ZHAO Wen-jiao,YAN Hong-wei,YANG Zhen,et al. An off-line optimization method for improving rolling force setting accuracy[J]. China Metallurgy,2017,27 (8):19.)

[4] 彭文,姬亚锋,陈小睿,等. 热轧非稳态过程轧制力自学习模型优化[J]. 东北大学学报(自然科学版),2019,40(10):1408. (PENG Wen,JI Ya-feng,CHEN Xiao-rui,et al. Optimization of rolling force Self-learning Model in hot rolling unsteady process[J]. Journal of Northeast University (Natural Science),2019,40 (10):1408.)

[5] 王振华,刘元铭,王涛,等. 粗轧过程中轧制力和宽展的预测与分析[J]. 钢铁,2022,57(9):95. (WANG Zhen-hua,LIU Yuan-ming,WANG Tao,et al. Prediction and analysis of rolling force and width spread in rough rolling[J]. Iron and Steel,2022,57(9):95.)

[6] 邱华东,田建艳,王书宇,等. 模糊神经网络融合建模方法及其在轧制力控制中的应用[J]. 中国冶金,2021,31(1):52. (QIU Hua-dong,TIAN Jian-yan,WANG Shu-yu,et al. Fuzzy neural network fusion modeling method and its application in rolling force control[J]. China Metallurgy,2021,31(1):52.)

[7] 陈治明,罗飞,黄晓红,等. 基于混沌优化支持向量机的轧制力预测[J]. 控制与决策,2009,24(6):808. (CHEN Zhi-ming,LUO Fei,HUANG Xiao-hong,et al. Rolling force prediction based on chaos optimization support vector machine[J]. Control and Decision Making,2009,24 (6):808.)

[8] 龚秀梅,王龙,高克伟,等. 极限学习机在中厚板轧制力预报中的应用[J]. 钢铁研究学报,2020,32(5):393. (JI Xiu-mei,WANG long,GAO Ke-wei, et al. Application of limit learning machine in rolling force prediction of medium and heavy plate[J]. Journal of Iron and Steel Research,2020,32 (5):393.)

[9] 刘明华,张强,刘英华,等. 基于机器学习的热轧轧制力预测[J]. 锻压技术,2021,46(10):233. (LIU Ming-hua,ZHANG Qiang,LIU Ying-hua,et al. Prediction of hot rolling force based on machine learning[J]. Forging Technology,2021,46 (10):233.)

[10] 马威,李维刚,赵云涛,等. 基于深度学习的热连轧轧制力预测[J]. 钢铁研究学报,2019,31(9):805. (MA Wei,LI Wei-gang,ZHAO Yun-tao,et al. Rolling force prediction of hot tandem rolling based on deep learning[J]. Journal of Iron and Steel Research, 2019, 31(9):805.)

[11] 徐冰冰,岑科廷,黄俊杰,等. 图卷积神经网络综述[J]. 计算机学报,2020,43(5):755. (XU Bing-bing,CEN Ke-ting,HUANG Jun-jie,et al. A review of graph convolutional neural networks[J]. Chinese Journal of Computers,2020,43(5):755.)

[12] Abu-El-Haija S,Kapoor A,Perozzi B,et al. N-GCN: Multi-scale graph convolution for semi-supervised node classification [C]//Proceedings of the 35th Uncertainty in Artificial Intelligence Conference. Tel-Aviv:Proceedings of Machine Learning Research,2020:841.