

基于增量学习树模型的带钢精轧宽度预测

王家亮¹, 王景成^{1,2}, 李继超¹

(1. 西安工业大学 电子信息工程学院, 陕西 西安 710012; 2. 上海交通大学 自动化系, 上海 200030)

摘要: 针对热轧带钢在轧制过程中出现的由于换辊或其他原因导致的精轧宽展预测模型不精准、热轧领域采集数据维度高和数据量大时训练时间长等问题, 提出了一种面向工况漂移的基于增量学习树模型的带钢精轧宽度预测方法。首先, 根据传统精轧机理模型得出变工况产生的原因和主要特征参数; 然后, 将在机理与数据层次上总结得到的有效数据特征作为预测模型的输入, 采用增量学习方法建立宽展预测模型; 最后, 通过实验设计, 在实际数据上将增量学习建立的树模型与其他机器学习模型在相同指标上进行对比。结果表明, 提出的基于增量学习的树模型在宽展预测中拥有较高的预测精度和较低的运行时间。

关键词: 热轧带钢; 精轧; 变工况; 自由宽展; 树的集成模型; 增量学习

DOI: 10.13330/j.issn.1000-3940.2024.02.019

中图分类号: TG33 **文献标志码:** A **文章编号:** 1000-3940 (2024) 02-0152-09

Prediction on strip finish rolling width based on incremental learning tree model

Wang Jialiang¹, Wang Jingcheng^{1,2}, Li Jichao¹

(1. School of Electronic and Information Engineering, Xi'an Technological University, Xi'an 710012, China;

2. Department of Automation, Shanghai Jiao Tong University, Shanghai 200030, China)

Abstract: Aiming at the problems of inaccurate prediction model of finish rolling width spreading due to roller change or other reasons during the rolling process of hot-rolled strip steel, and long training time when the dimension of data collected was high and the amount of data was large in the field of hot rolling, a width prediction method based on incremental learning tree model for operating condition drift was proposed. First, the causes and main characteristic parameters of the variable working conditions were obtained based on the traditional finish rolling mechanism model. Then, the effective data features summarized at the mechanism and data levels were combined as the input of the prediction model, and the width spreading prediction model was established by the incremental learning method. Finally, the tree model established by incremental learning was compared with other machine learning models with the same indicators on actual data by experimental design. The results show that the proposed tree model based on incremental learning has higher prediction accuracy and low running time in wide-spread prediction.

Key words: hot rolled strip; finish rolling; changing working conditions; free spread; tree integration model; incremental learning

随着“十四五”规划与 2035 年远景目标的提出, 国家正加速提升和突破基础工业以及产业技术等瓶颈与短板。而热轧带钢作为基础工业和流程制造工业等的重要原材料之一, 其质量与轧制过程中的宽度精度密切相关^[1]。

在实际生产中, 轧制过程普遍存在稳定性低且

精度控制水平低等问题, 从而出现带钢头部宽度超差、颈缩区宽度超差、尾部宽度超差和带钢整体宽度超差 4 大类, 严重影响了带钢的质量, 是较难控制的性能指标^[2-3]。常见的宽度控制策略主要存在于粗轧机组段, 包括短行程控制 (Short Stroke Control, SSC)^[4]、反馈宽度自动控制 (Feedback Automatic Width Control, AFB-AWC)^[5] 和前馈宽度控制 (Progressive Width Control, PWC) 等^[6]。在此基础上, 近年来, 为了满足高端板材日趋严苛的宽度质量要求, 国内外研究人员主要采用技术集成等控制策略^[7]。Cao J G 等^[8]提出了 VCR 变接触支持辊、ASR 非对称自补偿轧制工作辊系列技术、热磨辊工艺、工作辊液压弯辊与 ASR 非对称自补偿窜模型, 并投入我国电工钢生产过程中, 显著提高了生产实

收稿日期: 2023-06-29; 修订日期: 2023-09-18

基金项目: 国家自然科学基金资助项目 (62273234); 国家重点研发计划项目 (2022YFE0123400); 陕西省重点研发计划项目 (2022GY-236)

作者简介: 王家亮 (1999-), 男, 硕士研究生

E-mail: wangjljmei@163.com

通信作者: 王景成 (1972-), 男, 博士, 教授

E-mail: jewang@sjtu.edu.cn

绩。He N H 等^[9]针对热轧过程中的粗轧与精轧机组提出了一种适用于热轧高强钢的板形控制方法,并在实际热轧机上得到应用,取得了较为优异的性能表现与经济效益。

宽度控制是一个复杂的多变量问题,宽度的提前预知是精准控制宽度的基本要求,李维钢等^[10]提出了一种随机森林-差分进化(RF-DE)算法的铸坯宽度优化控制模型,以随机森林宽度预测模型的预测值与铸坯目标宽度的绝对误差为目标函数,构建带工艺约束条件的宽度优化控制模型,并使用差分进化算法求解得到连铸二冷工艺参数设定值,应用到实际生产后有效提升了铸坯宽度的合格率。而针对带钢精轧环节的宽度控制,由于精轧的宽度变化量较小,无法实行高效的宽度控制。传统的带钢宽度控制手段主要集中在粗轧段,但粗轧的宽度预设目标需要提前考虑粗轧入口目标宽度与精轧宽度变化量的总和。因此,精轧宽展的准确预测与带钢轧制过程中的宽度控制密切相关,对于保证热轧带钢的整体质量有着重要意义^[11]。

在实际生产过程中,热轧带钢的精轧现场环境具有高温、高湿、干扰性强的特点,任何操作环境的变化均会导致轧制工况漂移^[12]。一条生产线通常会在不同钢种、不同宽度和厚度层别的带钢上进行无周期性的加工作业,造成采集的数据变为变工况状态下的轧制数据。此外,轧制工况还易受到连续轧制的影响,造成轧辊的随机性磨损、生产过程中带钢原材料的性能波动和轧制参数微调等^[13]。在热轧领域,传统批量学习构建的神经网络方法通常将采集的数据视作同一工况来进行预测工作,对内部机理变化无法做到精确感知。当使用以往训练的模型对新轧制的带钢数据进行训练预测时,会出现效果显著变差的现象^[14]。

增量学习旨在应对机器学习中的概念漂移,既使得模型能够不断更新来学习新类别与新任务,同时又能够保持在已学习类别或任务上的性能的有效方法。孙世旭等^[15]针对刀具破损样本数量少且分布随工况变化的问题,提出一种半监督的增量学习方法,有效提升了对于不同工况下获取样本的刀具破损的识别效果。张煜莹等^[16]提出了一种基于增量学习的深度卷积诊断模型,通过增量学习实现知识迁移,学习新知识后复用到新数据集上,得到了良好的效果。任进等^[17]通过使用增量学习不断预测信道运行状态,通过学习不断更新 LSTM 神经网络的权重,实验表明,模型可以很好地适应无线信道的动

态变化且有效改善了时变信道的预测准确率。

上述文献通过实际工程论证,被证实相较于传统批量学习预测,增量学习方法可以有效改善批量建立的神经网络的预测效果。但同时,上述文献均建立在明确何值会发生变化的确定情况且拥有历史数据的基础上,通过重新打上标签进行增量训练从而完成预测或分类工作。而轧制是一个时间长、轧制钢种类型众多的过程,暂且不论是否能够一次性考虑到所有的钢种情况,若能够考虑到全部情况,其数据集将十分庞大且学习时间成本过高;此外,若进行分批次训练,则会导致后期训练数据取代前期训练的数据,出现神经网络的灾难性遗忘现象。而现行的带钢宽展预测仍大多采用批量方式构建模型^[18-19],无法对上述提出问题作出可行解。

因此,本研究根据上述总结的实际生产问题,完成以下工作:

(1) 从机理与数据层面分析内部因素发生变化的原因,总结与出口宽度关系密切的相关变量,有效降低数据维度;

(2) 在热轧领域提出基于增量学习的带钢宽展预测方法来改善处于变工况环境下的宽度预测准确率不佳的问题;

(3) 针对热轧场景生产数据多和轧制速度快等特征,引入近年热门的基于树的集成学习方法来进行带钢精轧的宽度预测,以有效提升预测模型的运行速度与预测准确度。

1 基于增量学习的宽展预测模型建立

1.1 精轧宽展机理及变工况分析

热轧带钢的轧制是一个多环节、多耦合的复杂控制过程,涉及到多种机械的相互配合,国内某钢厂 2050 mm 带钢热轧流程如图 1 所示,其主要由加热炉、侧压机、粗轧机组、精轧机组、层流冷却、输送辊道以及卷取机组构成。精轧机组一般由 7 个机架四辊轧机构成,包含 1 对支撑辊和 1 对工作辊。在带钢轧薄过程中,带钢在发生纵向延伸的同时不可避免地发生横向延伸,即精轧的自然宽展。传统数学模型对于精轧的自然宽展是建立在相关经验公式与参数的基础上,带钢精轧过程中的机理公式^[20]如式(1)~式(4)所示。

$$\Delta H_j = W_j \times [\exp(-0.5 \times \eta_j/100) - 1] \quad (1)$$

$$\eta_j = \exp[\alpha_1 - \alpha_2/(T + 273)] \times \theta_j^{\alpha_3} \times t_j^{\alpha_4} \quad (2)$$

$$t_j = D_j / [V_j(1 + R \times \lambda_j/k_j)] \quad (3)$$

$$W = \sum_{j=1}^7 \Delta H_j + C_j \quad (4)$$

式中: W 为总的精轧出口自然宽展偏差值, mm; ΔH_j 为第 j 个机架的自然宽展量偏差值, mm, $j=1, 2, \dots, 7$; $\alpha_1 \sim \alpha_4$ 为模型待定系数, 与轧制钢种的实际物理量相对应; t_j 为第 j 个机架间的通过时间量, s; η_j 为第 j 个机架轧制后的应变值; W_j 为第 j 个机架的设定出口宽度, mm; T 为机架间的轧制温度, K; θ_j 为第 j 个机架的轧辊设定的张力反馈, 取 10 MPa; D_j 为机架的相隔距离长度, m; R 为机架轧辊半径, m; λ_j 为第 j 个机架间轧辊中性角, ($^\circ$); k_j 为第 j 个机架设定的厚度量, mm; V_j 为机架间的轧制速度, ($\text{m} \cdot \text{s}^{-1}$); C_j 为针对第 j 个机架轧制的补偿量, mm。

带钢生产线由于其生产系统庞大且设备机理结构复杂的特点, 随着服役时间的增长, 部分内部零件不可避免地发生退化, 导致内部潜在故障

点较多。而进行钢材轧制时, 高速、高压与高湿的内部环境, 为即时检测带来了很大的挑战^[21]。针对精轧宽展过程, 仅在入口与出口处设有测宽仪, 内部环境没有准确的传感器采集数据, 无法做到对变化的精准感知, 单从轧辊的情况来说, 彭文等^[22]使用 ABAQUS 建立了轧制过程中的轧辊-工件热力耦合的有限元分析模型, 详细分析了不同工况情况下, 轧辊的磨损产生状况, 并实现了轧制过程中的工作辊磨损的高精度预测, 有力佐证了变工况状态下轧辊的变化会给带钢的板形控制带来不可忽视的影响。而轧制一经开始则需要完成大量轧制任务, 在进行各类轧制任务的过程中, 轧辊难免会受到磨损, 出现工况漂移的情况。由式 (3) 可知, 轧辊的状态与精轧宽展密切相关, 但由于热轧过程中高温、高湿, 轧辊的状态无法进行实时监测, 在执行换辊操作前, 采集的数据则一直为原始标称值, 导致了轧制工况变化而无法及时感知, 造成预测效果差。

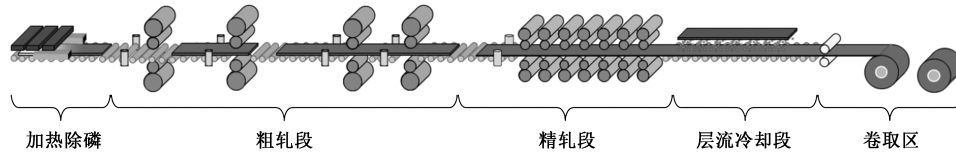


图 1 国内某钢厂 2050 mm 带钢热轧流程图

Fig. 1 Hot rolling flow diagram of 2050 mm strip in a domestic steel mill

1.2 基于树的集成学习预测模型

1.2.1 随机森林 (RF) 和梯度提升树 (XGBoost)

在集成学习中, 主要分为套袋法 (Bagging) 和提升法 (Boosting), 随机森林 (Random Forest, RF) 作为一种广泛使用的 Bagging 类型的多元分类器, 每个随机森林树都是在不同的采样训练集上独立进行训练的, 而由于每个子树可以同时构建, 因此, 相较于传统机器学习算法, 大大提升了训练时长。此外, 它还对多数决策树以投票的方法平均最终结果, 从而减少过度拟合问题, 显著提升预测性能, 其中随机森林函数 $\psi(x)$ 可以由式 (5) 来表示。

$$\psi(x) = \frac{1}{N_{\text{Trees}}} \sum_{u=0}^{N_{\text{Trees}}} T_u(x) \quad (5)$$

式中: N_{Trees} 为树的数量; $T_u(x)$ 为随机生成的第 u 个树, $u=0, 1, \dots, N_{\text{Trees}}$ 。

极限梯度提升树 (Extreme Gradient Boosting, XGBoost) 是一种较为受欢迎并被广泛使用的算法。受限于梯度决策树 (Gradient Boosting Decision Tree, GBDT) 的高额训练时间复杂度, Chen T Q 等^[23]开

发了一种基于 GBDT 算法的改进的 XGBoost 算法, 在 GBDT 高精度预测的基础上, 通过迭代整合多个决策树模型, 添加正则化项来抑制模型过拟合, 同时利用最小化的损失函数来训练模型的预测能力, 从而实现时间和性能更优的预测模型。其正则化目标函数 κ 可由式 (6) 表示。

$$\kappa = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (6)$$

式中: i 为样本序列号; y_i 与 \hat{y}_i 分别为第 i 个样本的真实值与预测值; l 为衡量预测值与真实值的损失函数; $\Omega(f_k)$ 为第 k 棵树的正则化项。

1.2.2 轻量级梯度提升算法 (LightGBM)

轻量级梯度提升机 (Light Gradient Boosting Machine, LightGBM) 是微软在 2017 年提出的 XGBoost 算法的改进版, 相较于前者, LightGBM 采用直方图算法将遍历形式从遍历数据转变为遍历直方图, 有效降低了时间复杂度; 对比 XGBoost 的 level-wise 的分裂策略, LightGBM 采用的 leaf-wise 分裂策略可以快速找出最大分裂增益的叶, 然后持续进行分裂, 相

较于 level-wise 的无差别分裂, 有效提升了训练速度。

1.3 基于增量学习的机器学习预测模型构建

机器学习拥有较强的非线性映射能力以及自学习、自适应能力, 可以广泛应用在不同的实际场景下。但是, 随着应用范围的扩大和数据量的增长, 随之带来的“灾难性遗忘”和训练时间长, 均限制了其精度的提升。

而上述问题通过增量学习方法可以有效解决。增量学习方法^[24]的定义是通过给定的数据流样本 $\{s_1, s_2, \dots, s_t\}$ 生成一系列的模型 $\{h_1, h_2, \dots, h_t\}$, 其表达式为:

$$s_t = (x_t, y_t) \in \mathbf{R}^e \times \{1, 2, \dots, C\} \quad (7)$$

$$h_t \in \mathbf{R}^e \times \{1, 2, \dots, C\} \quad (8)$$

式中: s_t 为 t 时刻产生的带有标签的训练样本; h_t 为 t 时刻的模型; h_{t-1} 仅由 $(t-1)$ 时刻的模型 h_{t-1} 和最近的 p 个训练样本 $\{s_{t-p+1}, s_{t-p+2}, \dots, s_t\}$ 得到; p 为增量学习样本的数量上限; x_t 和 y_t 分别为 t 时刻的训练特征样本和目标值; C 为目标变量的数量; \mathbf{R} 为数据空间; e 为数据空间的维度。

常见的增量学习种类分别是基于正则化和基于知识回放的方式。基于正则化进行增量学习的思想是通过给新任务加入损失函数约束来保护旧的知识不被破坏, 从改善网络结构, 其一定程度上缓解了灾难性遗忘的问题; 而基于知识回放的增量学习方式的主要思想为“温故而知新”, 即将旧数据与新数据一起训练模型, 从而提升对过往数据的预测效果^[25-26]。鉴于钢铁行业数据体量大、轧制特征维度高的特点, 知识回放的方法势必会导致训练时间过长, 需要较高的时间成本。因此, 本研究选择基于正则化的方法进行模型构建。

虽然, 轧制过程中会出现不同型号、不同层别的钢种, 导致采集的数据集会不太相同, 但是其本质还是对于带钢的宽展进行预测, 是归类于相同任务, 因此, 本研究采用的训练策略为: 假设现有数据集 A 与 B , 在进行增量训练时, 在数据集 A 训练完成的模型基础上, 不增加新的决策树 (网络节点), 保持原结构不变, 重新计算叶子节点权重, 开始对数据集 B 进行训练, 当出现欠拟合、过拟合状态时也可适当增加新的决策树 (网络节点)。

当完成相差不大的任务数据集 A 与 B 时, 模型仅需要在训练时调整相关树的权重, 使其适应整体特点, 使处理相似轧制过程时其均拥有较好的结果。

其整体算法的实现流程如图2所示。

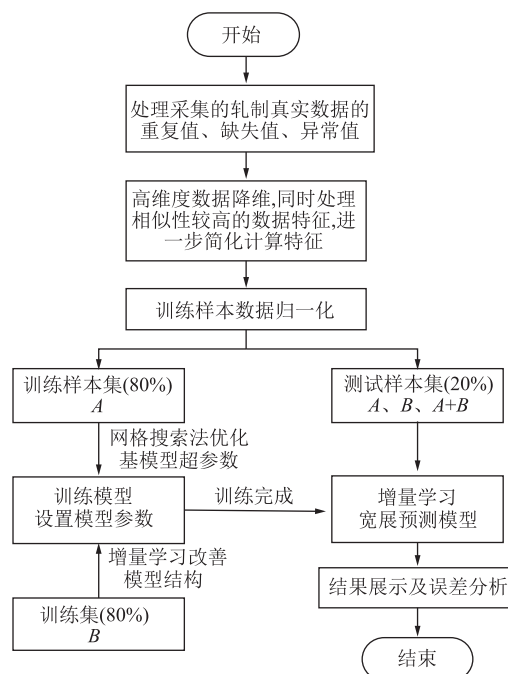


图2 增量学习宽展预测模型实现流程图

Fig. 2 Implementation flow chart of width prediction model based on incremental learning

(1) 使用皮尔逊相关性分析方法对采集的数据进行降维处理, 减少训练数据的特征数量, 达到加快训练速度的目的;

(2) 将采集数据按照宽度层别分为 A 与 B 两大类, 并按 8 : 2 将其进行划分, 同时构建基础预测模型, 导入数据集 A 进行训练。此外, 采用网格搜索法对模型最优化参数进行调整, 达到较优的预测性能;

(3) 在步骤 (2) 模型的基础上采用增量学习方式对数据集 B 模型进行训练, 更新原始模型节点权重以及网络结构, 得到包含数据集 A 与数据集 B 特征的预测模型;

(4) 分别代入数据集 A 、数据集 B 以及数据集 $A+B$ 的 20% 作为测试集进行预测, 并对所得结果进行量化评价, 分析模型的性能结果。

2 实验设计与结果分析

2.1 数据的预处理与特征工程

实验数据来自于国内某钢厂的 2050 mm 带钢生产线的实际生产数据, 从生产线的各个工业传感器共计 79 个特征因素和 29478 条轧制数据。其中包含了各个轧制机架的相关物理系数、轧制力、轧制速

度、轧制温度以及轧制钢种的化学成分等相关数据。在导入训练模型之前,需要对这些数据进行预处理。主要步骤包括:删除重复值、补全或删除缺失值以及处理异常值与孤立点。针对重复值可以直接使用相同值筛选进行剔除;针对缺失值则使用零值或空值检测进行挑选并删除;而针对异常值则可以使用基于正态分布的 3σ 准则来将异常值进行筛选并删除,至此完成数据预处理工作。

通过上述数据清洗方法,发现不存在重复值,但存在缺失值与异常值,共 1680 条数据,得到剩余 27798 条符合要求的数据。随后将数据均分为 A 与 B 两个数据集用于接下来的模型训练。

而在采集的数据中,影响轧制的特征众多,经过统计,共有 75 个维度,为了保证训练速度与精度,需要选取高质量与相关性强的特征因素进行模型训练。从 1.1 节的机理分析可知:可大致将影响宽度变化的轧制因素分为物理与几何因素。物理因素即摩擦因数、金属化学成分、轧制速度等,其会影响变形区内部的作用力,尤其是摩擦力等关键因素;几何因素则是指产生于变形区内的高度、长度、宽度、轧辊半径以及变形区的横截面形状等关键几何特征。因此,在精轧自然宽展模型中,基于机理分析,可以总结出决定带钢宽度的关键输入变量为钢族、出口厚度和宽度以及轧辊压下 4 大类变量。

继而从数据层面,皮尔逊相关性系数可以很好地描述多变量之间的相关性。其主要原理为:假设存在两个样本集 $X: \{x_1, x_2, \dots, x_n\}$ 和 $Y: \{y_1, y_2, \dots, y_n\}$, 则皮尔逊系数 $Corr$ 可以用式 (9) 进行计算。

$$Corr = \frac{\sum_{i=1}^n \frac{x_i - E(X)}{\sigma_X} \cdot \frac{y_i - E(Y)}{\sigma_Y}}{n} \quad (9)$$

式中: $E(X)$ 和 $E(Y)$ 分别为样本集 X 和 Y 的均值; x_i 和 y_i 分别为样本集 X 和 Y 中的元素; σ_X 和 σ_Y 分别为样本集 X 和 Y 的方差; n 为样本总数。

将所有特征采用皮尔逊相关系数进行计算,同时,将与真实自然宽展相关性低于 0.1 的相关性特征舍去。同时,考虑到采集数据是根据一定时刻采集的生产数据来进行分析的,存在偶然性与不均匀性。因此,本文结合机理层面与数据层面的特征变量分析,从而总结得出所选的数据特征,具体所选取的特征数据如表 1 所示。其中, F1~F7 表示机架 1~机架 7。

表 1 轧制的关键特征

Table 1 Key features of rolling

编号	参数	单位
1	碳当量	—
2~3	宽度层别与厚度层别	—
4	精轧入口宽度	mm
5~12	机架之间的辊缝 (F1~F7)	mm
13~20	轧制速度 (F1~F7)	($m \cdot s^{-1}$)
21~28	轧制温度 (F1~F7)	℃
29~36	各轧辊轧制力 (F1~F7)	kN
37	精轧出口宽度	mm

2.2 实验环境与评价指标

本文提出的精轧自由宽展预测模型算法采用 Python3.10 作为编程语言,软件开发采用 Pytorch 框架,调用 Pandas、Numpy 和 Sklearn 集成库。为了验证增量学习宽展预测模型的预测性能,采用相同结构的随机森林回归模型、XGBoost 模型、LightGBM 模型,作为对比,加入 BP 神经网络以及相对应的增量学习 BP 网络,在相同数据集下展开对比实验,其中,模型的网络结构以及超参数设定如表 2 所示。仿真环境为 Intel i5-7300HQ、RAM 12 GB 和 128 GB SSD。

表 2 模型超参数设置

Table 2 Hyperparameter setting of model

预测模型	模型结构以及参数设置
BP/IL-BP	'hidden_layer_size': (95, 43, 21), 'activation': 'relu', 'tol': 1e-6, 'learning_rate_init': 0.0003
RF/IL-RF	'max_depth': 7, 'min_samples_leaf': 2, 'max_features': 6, 'n_estimators': 20
XGBoost/ IL-XGBoost	'booster': 'gbtree', 'gamma': 1.1, 'max_depth': 15, 'lambda': 1, 'subsample': 0.7, 'eta': 0.04,
LightGBM/ IL-LightGBM	'boosting_type': 'gbdt', 'max_depth': 2, 'num_leaves': 31, 'learning_rate': 0.01,

将本数据集按 8:2 进行划分,训练特征因素为表 1 中经过相关性分析所筛选的特征,输入目标变量为精轧出口自由宽展值,以真实采集的自由宽展值作为基准值 M ,模型预测值则为 $M + \Delta M$ 。同时,由于轧制现场需要预测模型具有一定的稳定性以及可迁移性。因此,本文对实验模型采用相同的评判标准来描述各个模型的真实值与预测值的偏移量。

采用平均绝对值误差（Mean Absolute Error, MAE ）与均方误差（Mean Squared Error, MSE ）来描述偏移量的大小与趋势；采用决定系数（R-Square, R^2 ）来描述模型回归曲线的拟合程度。具体计算公式如式（10）~式（12）所示。

$$MAE = \frac{1}{m} \sum_{i=1}^m |(y_i - \hat{y}_i)|$$

(10)

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2$$

(11)

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y}_i)^2}$$

(12)

式中： m 为样本总量； y_i 、 \hat{y}_i 、 \bar{y}_i 分别为第 i 个样本的自然宽展的真实值、预测值和真实值的平均值， $i=1, \dots, m$ 。

3 实验结果分析

3.1 基于增量学习的宽展预测结果分析

将 4 个模型的测试集上的预测值与相应的实际值进行比较，图 3 为 4 种通过批量学习建立的基模型与其对应的使用增量学习建立的模型在测试集 $A+B$ 上的交叉验证 5 次的 R^2 结果对比图。由图 3 可知，经过增量学习建立的模型相较于批量学习建立的精度均有所提升。

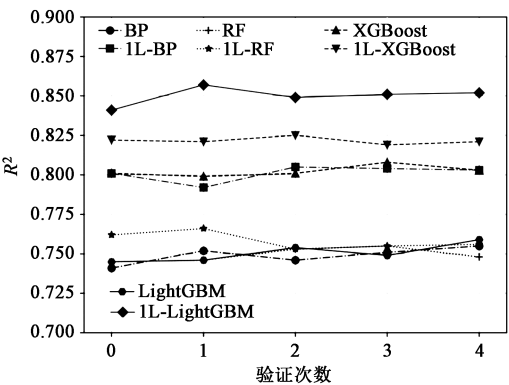


图 3 各模型基模型与增量模型预测精度比较
Fig. 3 Comparison of prediction accuracy between base models and incremental models for each models

图 4 仔细比较了各个基模型、各个基模型对应的增量模型与真实值的拟合程度，由图 4 可知，相较于批量学习的方式，加入增量学习的模型在变工况情况下对于精轧宽展预测的波动性较小，拟合程度高。

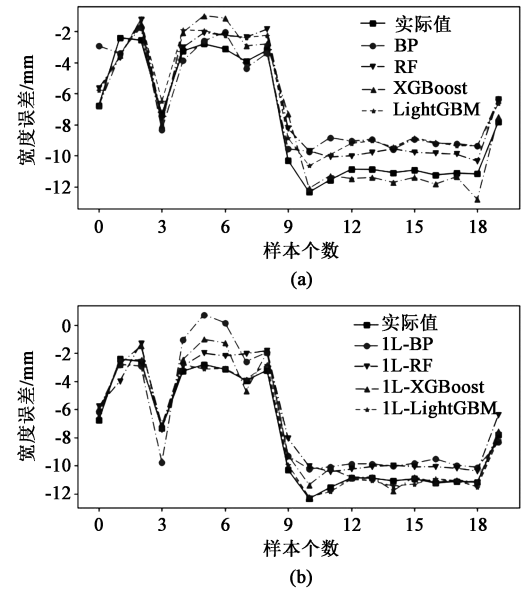


图 4 批量模型 (a) 和增量模型 (b) 在数据集 $A+B$ 上的预测结果
Fig. 4 Prediction results of batch model (a) and incremental model (b) in dataset $A+B$

为进一步表现增量学习对于神经网络的灾难性遗忘问题的解决问题，表 3 为各模型的预测指标，其中，加粗数据为各个测试集下的最佳性能数据。在数据集 A 的表现上，使用了增量学习的 XGboost 算法取得了最好的效果，分析原因可能是在学习过程中，IL-BP、IL-RF、IL-LightGBM 模型在为了平衡数据集 B 的过程中，改变了先前训练数据集 A 的有关叶子节点的权重，从而导致了预测精度有所下降。

表 3 模型在不同数据集上的性能对比						
Table 3 Performance comparison of models in different datasets						
模型	测试集 A		测试集 B		测试集 A+B	
	MAE	MSE	MAE	MSE	MAE	MSE
BP	1.293	2.471	1.268	2.623	1.283	2.548
IL-BP	1.159	2.126	1.247	2.781	1.201	2.454
RF	1.055	1.807	1.343	2.673	1.198	2.241
IL-RF	1.026	1.771	1.323	2.609	1.174	2.191
XGBoost	1.216	2.301	1.307	2.501	1.262	2.401
IL-XGboost	1.017	1.671	1.277	2.443	1.147	2.057
LightGBM	1.043	1.711	1.367	2.754	1.205	2.232
IL-LightGBM	1.079	1.704	1.209	2.061	1.144	1.882

而在数据集 B 和数据集 $A+B$ 中，无论是单独对于数据集 B 进行预测，还是对于 $A+B$ 整体数据集进行预测，可以看出，使用增量学习模型的预测效果

均高于未使用增量学习时的预测效果。这是因为：在模型训练过程中，通过增量学习使得原本模型在适应数据集 *A* 的基础上，通过调整叶子节点权重或者增加新的叶子节点对数据集 *A* 与数据集 *B* 的特征进行了平衡，使两个数据集上均拥有不错的预测效果。从而验证了通过增量学习方式建立模型，可在一定程度上缓解因为重新训练而导致的神经网络的遗忘问题。

3.2 基于增量树模型的宽展预测结果分析

根据表 3，可初步看出基于树的集成学习模型及其对应的增量学习模型相较于 BP 预测模型和 IL-BP 预测模型在预测指标 *MAE* 与 *MSE* 上均取得了较好的结果。

为了进一步证实树模型在处理大量轧制数据时具有一定的精度度与较低的时间消耗，本研究采用常见的几种预测算法，如：KNN、SVR、BP，和常见的基于树的算法，如：GDBT、RF、AdaBoost、XGBoost、LightGBM，将在 1000、10000、100000 条数据量的混合轧制数据上的运行时间与准确度来作对比分析，此外，模型继续使用表 2 所示超参数。其他模型采用自带 Sklearn 库中的标准基模型参数，实验结果如表 4 所示。

表 4 不同模型在预测精度和运行时间上的性能对比

Table 4 Performance comparison of different models in terms of prediction accuracy and runtime

模型	数据量/条					
	1000		10000		100000	
	MAE	时间/s	MAE	时间/s	MAE	时间/s
Mechanism	1.57	—	1.52	—	1.46	—
Linear	1297	0.010	2.36	0.04	1.84	0.72
KNN	1.43	0.160	1.81	0.03	1.74	2.29
Ridge	1.94	0.003	2.32	0.01	1.84	0.08
Lasso	1.62	0.002	1.43	0.01	1.85	0.13
其他 BP	1.68	7.380	1.61	178.18	1.39	589.61
模型 SVR	1.73	0.09	1.55	8.17	1.51	1086.10
DecisionTree	3.19	0.03	2.74	0.51	2.52	6.82
ExtraTree	2.01	0.01	2.90	0.12	2.62	1.81
Adaboost	1.74	0.38	2.10	3.91	2.11	60.65
Bagging	2.12	0.18	1.67	2.83	1.74	42.24
GDBT	2.13	1.88	1.53	12.01	1.48	112.70
增量 RF	1.64	1.57	1.18	8.01	1.53	339.95
学习 XGBoost	1.75	0.24	1.49	6.14	1.50	14.18
模型 LightGBM	1.67	0.22	1.25	1.37	1.33	4.66

表 4 中可以看出，批量建立的回归预测模型，如 Linear 线性模型、KNN、Ridge 回归、Lasso 回归模型，可以达到较快的计算速度，但在计算的精确度上相较于原始产线机理计算模型 Mechanism 还有待提升，究其原因还是因为针对轧制情况的变工况场景无法做到精准识别，当数据量上升时，后续数据会逐渐覆盖前面的经验，也导致了预测效果不佳；而 BP 网络和 SVR 预测模型虽然达到了较好的精度，但需要大量的训练时间，无法满足实际生产的实时性；而基于树的集成学习模型，相较于前面所提出的常见的机器学习预测算法，在保持误差指标较为优秀的情况下，达到了较低的运行时间，其中基于 LightGBM 算法，在使用 100000 条仿真数据的时候，实现了较好的预测误差的同时，也达到了较短的运行时间，证明了其在面对数据维度高和海量数据的轧钢场景的可靠运用效果。

考虑到有一些模型现阶段无法进行增量学习训练，因此，对于增量学习预测精度度与模型计算时间，采用本文此前提出的 BP、RF、XGBoost 和 LightGBM 模型进行对比。表 5 比较了通过增量学习建立模型的情况下各模型的网络在 10000 条数据量下的训练时间以及预测精度度。

表 5 增量学习模型预测误差与运行时间对比

Table 5 Comparison of prediction error and runtime for incremental learning models

模型指标	MAE	时间/s
IL-BP	1.518	27.18
IL-RF	1.178	7.72
IL-XGBoost	1.190	1.44
IL-LightGBM	1.170	1.14

通过进一步对比可知，采用增量学习的方式建立模型，基于树的集成学习预测模型相较于 BP 神经网络有效提升了误差精度 *MAE* 和缩短了运行时间。其中，IL-LightGBM 模型能够在较短的时间内，更为精确地进行预测工作。因此，可以认为基于树的集成学习模型具有更快的运行速度，以及更好的泛化性能，可以有效应用在轧制过程中的精轧宽度预测，起到一定的参考作用。

4 结 论

(1) 通过分析实际轧制过程中出现的变工况场

景, 同时发现传统批量构建模型的方式无法解决当前变工况所带来的问题, 从而提出了基于增量学习的宽展预测模型, 最终通过实际数据实验仿真, 有效解决了当前面临的变工况环境造成预测不准的问题。

(2) 引入近年热门的基于树的集成学习方法来进行带钢精轧的宽度预测。通过建立批量学习预测模型与增量学习预测模型, 在相同数据集下, 对预测误差与模型运行速度进行对比, 总结出基于树的集成模型在生产数据多和轧制速度快等热轧场景的可行性。同时, 快速与精确的宽度预测结果, 还为实现建立在精准预测基础上的高精度宽度控制方法提供了有效依据, 为实际生产的在线应用提供了一种新思路。

参考文献:

- [1] 何安瑞. 热轧宽带钢板形控制技术的现状及未来发展 [J]. 轧钢, 2022, 39 (3): 1-10.
He A R. Present situation and future development of profile and flatness control technologies of hot rolled wide strip [J]. Steel Rolling, 2022, 39 (3): 1-10.
- [2] 黄爽. 热轧带钢宽度超差控制技术的改进与应用 [D]. 沈阳: 东北大学, 2015.
Huang S. Improvement and Application of Width Tolerance Control Technology in Hot Strip [D]. Shenyang: Northeastern University, 2015.
- [3] 张宏献, 回士敏, 马欣然. 提高热轧带钢成品的宽度控制精度 [J]. 中国冶金, 2020, 30 (8): 60-63.
Zhang H X, Hui S M, Ma X R. Improvements of width controlling precision for finished hot-rolled strips [J]. China Metallurgy, 2020, 30 (8): 60-63.
- [4] 彭文, 马更生, 曹剑钊, 等. 热轧带钢短行程控制自适应策略 [J]. 东北大学学报: 自然科学版, 2016, 37 (3): 343-346.
Peng W, Ma G S, Cao J Z, et al. Adaptive strategy of short stroke control in tandem hot rolling [J]. Journal of Northeastern University: Natural Science, 2016, 37 (3): 343-346.
- [5] 邹俊, 傅新, 杨华勇, 等. 粗轧基础自动化宽度控制和仿真系统研究 [J]. 钢铁, 2005, (4): 49-52.
Zou J, Fu X, Yang H Y, et al. Simulation system investigation and basic automation of width control for rough rolling [J]. Iron and Steel, 2005, (4): 49-52.
- [6] 陈丰, 杨子江, 王庆军, 等. 热连轧带钢生产线计算机控制系统的研发与应用 [J]. 轧钢, 2019, 36 (3): 59-65.
Chen F, Yang Z J, Wang Q J, et al. Development and application of control system of hot continuous strip rolling line [J]. Steel Rolling, 2019, 36 (3): 59-65.
- [7] 曹建国, 宋纯宁, 王雷雷, 等. 新一代高技术轧机电工钢矩形断面板形控制创新研究 [A]. 中国金属学会. 第十三届中国钢铁年会论文集: 4. 轧制与热处理 [C]. 北京: 冶金工业出版社, 2022.
- [8] Cao J G, Wei G C, Zhang J, et al. VCR and ASR technology for profile and flatness control in hot strip mills [J]. Journal of Central South University of Technology, 2008, 15: 264-270.
- [9] He N H, Wang C X, Yang Q. High-precision section control technology for high-strength yoke steel strip [J]. Materials Science Forum, 2019, 4734: 212-221.
- [10] 李维刚, 淳李良, 李阳, 等. 不锈钢铸坯宽度优化控制模型研究及应用 [J]. 钢铁, 2021, 56 (12): 75-84.
Li W G, Chun L L, Li Y, et al. Research and application of optimal control model for slab width of stainless steel [J]. Iron and Steel, 2021, 56 (12): 75-84.
- [11] Ji Y J, Liu S X, Zhou M C, et al. A machine learning and genetic algorithm-based method for predicting width deviation of hot-rolled strip in steel production systems [J]. Information Sciences, 2022, 589: 360-375.
- [12] 陈鹏宇, 隆建, 杨明磊, 等. 基于工业数据的溶剂脱沥青装置多工况建模 [J]. 控制工程, 2020, 27 (11): 2002-2009.
Chen P Y, Long J, Yang M L, et al. Multi-conditional modeling of solvent deasphalting device based on industrial data [J]. Control Engineering of China, 2020, 27 (11): 2002-2009.
- [13] 张瑞成, 曹志新. 基于 EEMD-LSTM 的冷连轧机振动预测研究 [J]. 锻压技术, 2022, 47 (9): 174-181.
Zhang R C, Cao Z X. Research on vibration prediction for tandem cold rolling mill based on EEMD-LSTM [J]. Forging & Stamping Technology, 2022, 47 (9): 174-181.
- [14] 董家欢, 邱清盈, 管成. 机械剩余使用寿命预测模型的增量学习方法 [J/OL]. 计算机集成制造系统: 1-15 [2023-05-27].
Dong J H, Qiu Q Y, Guan C. Incremental learning methods for remaining useful life prediction models of machinery [J/OL]. Computer Integrated Manufacturing Systems: 1-15 [2023-05-27].
- [15] 孙世旭, 胡小锋, 夏铭远. 变工况刀具破损监测的半监督增量学习方法 [J]. 计算机集成制造系统, 2022, 28 (9): 2805-2814.
Sun S X, Hu X F, Xia M Y. Semi-supervised incremental-learning method for tool breakage detection under variable operation conditions [J]. Computer Integrated Manufacturing Systems, 2022, 28 (9): 2805-2814.
- [16] 张煜莹, 陆艺, 赵静. 基于增量学习的数控机床故障诊断系统 [J]. 计量学报, 2022, 43 (11): 1456-1463.
Zhang Y Y, Lu Y, Zhao J. Fault diagnosis system of numerical control machine based on incremental learning [J]. Acta Metrologica Sinica, 2022, 43 (11): 1456-1463.
- [17] 任进, 邵淑颖, 何怡怡. 基于增量学习的时变信道预测方法

- [J]. 无线电工程, 2023, 53 (4): 815-823.
- Ren J, Shao S Y, He Y Y. Time-varying channel prediction method based on incremental learning [J]. Radio Engineering, 2023, 53 (4): 815-823.
- [18] 彭文, 陈庆安, 马更生, 等. 热连轧宽度自适应模型优化 [J]. 东北大学学报: 自然科学版, 2017, 38 (9): 1243-1246.
- Peng W, Chen Q A, Ma G S, et al. Optimization of the width adaption model in hot tandem rolling [J]. Journal of Northeastern University: Natural Science, 2017, 38 (9): 1243-1246.
- [19] 王晓雯, 张勇军, 郭强, 等. 基于轧制机理和混合神经网络的热轧精轧带宽预测 [J]. 中国冶金, 2023, 33 (2): 114-120.
- Wang X W, Zhang Y J, Guo Q, et al. Prediction of hot strip width based on rolling mechanism and hybrid neural network [J]. China Metallurgy, 2023, 33 (2): 114-120.
- [20] 张峥, 仲兆准, 李阳, 等. 基于深度学习的带钢精轧过程自由宽展预测 [J]. 中国冶金, 2022, 32 (11): 121-127.
- Zhang Z, Zhong Z Z, Li Y, et al. Prediction of lateral spread for hot strip finishing mill based on deep learning [J]. China Metallurgy, 2022, 32 (11): 121-127.
- [21] 宋纯宁, 曹建国, 王雷雷, 等. 六辊冷连轧机电工钢矩形断面控制弯辊力模型 [J]. 哈尔滨工业大学学报, 2022, 54 (7): 143-150.
- Song C N, Cao J G, Wang L L, et al. Model of rectangular section control roll bending force for electrical steel in six-high tandem cold mill [J]. Journal of Harbin Institute of Technology, 2022, 54 (7): 143-150.
- [22] 彭文, 孙佳楠, 李旭东, 等. 板带热轧过程工作辊磨损预测研究 [J]. 塑性工程学报, 2023, 30 (5): 214-225.
- Peng W, Sun J N, Li X D, et al. Study on prediction of work roll wear during hot strip rolling [J]. Journal of Plasticity Engineering, 2023, 30 (5): 214-225.
- [23] Chen T Q, Guestrin C. XGBoost: A scalable tree boosting system [A]. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [C]. ACM, 2016.
- [24] Losing V, Hammer B, Wersing H. Incremental on-line learning: A review and comparison of state of the art algorithms [J]. Neurocomputing, 2018, 275: 1261-1274.
- [25] 朱飞, 张煦尧, 刘成林. 类别增量学习研究进展和性能评价 [J]. 自动化学报, 2023, 49 (3): 635-660.
- Zhu F, Zhang X Y, Liu C L. Class Incremental learning: A review and performance evaluation [J]. Acta Automatica Sinica, 2023, 49 (3): 635-660.
- [26] 罗杨洋, 韩锡斌. 基于增量学习算法的混合课程学生成绩预测模型研究 [J]. 电化教育研究, 2021, 42 (7): 83-90.
- Luo Y Y, Han X B. A model for predicting student performance in hybrid courses based on incremental learning algorithm [J]. e-Education Research, 2021, 42 (7): 83-90.

2024 年发布的 4 项锻压领域国家和行业标准内容简介

GB/T 43479—2023 金属旋压成形性能与试验方法 成形性能、成形指标及通用试验规程

本文件规定了金属材料室温旋压成形的成形性能、成形指标, 并描述了性能试验的通用性操作步骤, 包括试样准备, 试验旋轮、芯模与顶头准备, 试验辅助装置与机床准备, 试样润滑与冷却, 预试验, 正式试验, 数据采集与计算, 以及试验报告。

本文件既适用于金属拉深旋压、缩径旋压、剪切旋压与流动旋压等旋压成形性能的检测, 还适用于金属非轴对称件缩径旋压与非圆横截面件旋压等特种旋压成形性能的检测。

JB/T 14450—2023 铝合金车轮摆动辗压-旋压复合成形件 通用技术规范

本文件界定了铝合金车轮摆动辗压-旋压复合成形件的术语和定义, 规定了技术要求, 描述了相应的试验方法, 规定了检验规则以及标志、包装、运输和贮存等要求。

本文件适用于商用车辆铝合金车轮摆动辗压-旋压复合成形件的制造, 乘用车铝合金车轮摆动辗压-旋压复合成形件的制造可参照使用。

JB/T 14451—2023 钢质锻件锻造生产能源消耗限额及评价方法

本文件规定了钢质锻件生产的单位能源消耗计算方法以及评价标准和评价方法。

本文件适用于以天然气或电为能源的钢质锻件生产过程中加热、锻造、热处理工序的能源消耗核算及评价活动。

JB/T 14452—2023 钢质楔横轧件材料消耗工艺定额编制要求

本文件规定了钢质楔横轧件材料消耗工艺定额的编制方法和明细表。

本文件适用于钢质楔横轧件的材料消耗工艺定额的编制活动。

全国锻压标准化技术委员会秘书处