

基于宽度学习系统的仓储粮情风险点预测模型

廉飞宇, 秦瑶, 付麦霞*

河南工业大学 信息科学与工程学院, 粮食信息处理与控制教育部重点实验室, 河南 郑州 450001

摘要:及时的仓储粮情预测是保证储粮安全的必要手段。目前,传统预测方法多从某一侧面对仓储粮情进行预测,无法实现对仓储粮情风险的精准综合评估,而深度学习方法则存在着所需训练样本数量巨大、训练难度高、时间长等瓶颈问题。针对这一现状,采用基于宽度学习的特征提取与融合方法,以及基于增量学习的训练方法(增强节点和输入数据增量算法),结合粮情数据的多模态特征,在宽度学习系统现有框架的基础上,提出了基于宽度学习系统的粮情风险预测模型。结果表明,与现有深度学习模型相比,在不降低预测准确度的前提下,预测模型大大节省了模型训练时间,降低了训练难度。预测模型成为深度学习模型的一种有效替代方案。

关键词:粮情风险点;宽度学习系统;多模态数据;增量学习;典型相关性分析

中图分类号:TS210

文献标志码:A

文章编号:1673-2383(2023)03-0104-09

DOI:10.16433/j.1673-2383.2023.03.014

Risk point prediction model of warehouse grain situation based on broad learning system

LIAN Feiyu, QIN Yao, FU Maixia*

College of Information Science and Engineering, Key Laboratory of Grain Information Processing and Control, Ministry of Education, Henan University of Technology, Zhengzhou 450001, China

Abstract: Timely forecast of stored grain condition is a necessary means to ensure the safety of stored grain. At present, traditional forecasting methods mostly predict the stored grain situation from one side, and cannot achieve accurate comprehensive assessment of the risk of stored grain situation. However, deep learning methods have bottleneck problems such as a large number of required training samples, high training difficulty and long training time. In view of this situation, by using the feature extraction and fusion method based on broad learning and the training method based on incremental learning, and combined with the multi-modal characteristics of grain situation data, a grain situation risk prediction model was proposed on the basis of the existing framework of the broad learning system. The results showed that, compared with the existing deep learning model, the training difficulty and time-consuming of the model were greatly reduced without reducing the accuracy of prediction. The predictive model proposed in the paper may be an effective alternative to deep learning models.

Key words: grain situation risk point; broad learning system; multi-modal data; incremental learning; canonical correlation analysis

粮食储备是保证我国粮食安全的重要举措。(简称粮情),还需要及时地预测仓储粮情的变化。为了保证储粮安全,除了要实时监测储粮状态趋势。仓储粮情风险点预测,就是要预测一个时

收稿日期:2022-07-17

基金项目:粮食信息处理与控制教育部重点实验室开放基金项目(KFJJ-2021-103);河南省高等学校重点科研项目(22A510014);河南省高校青年骨干教师培育项目(2020GGJS084);河南工业大学青年骨干教师培育项目

作者简介:廉飞宇(1970—),男,河南郑州人,副教授,研究方向为智能信息处理。

*通信作者:付麦霞,副教授,E-mail:fumaixia@126.com。

间点,储粮到了这个时间点,极有可能发生霉变和品质劣变而导致不可食用,必须进行人工处理。仓储粮情风险点的预测准确度与多种因素有关。相关研究表明,仓储粮情预测是一个十分复杂的问题^[1],也是目前粮食储藏领域迫切需要解决的关键科技问题。

对于预测问题,传统机器学习的方法有线性回归(LR)、支持向量回归(SVR)^[2]、随机森林^[3]、ARIMA^[4]、灰色预测模型^[5]、BPNN^[6]等,已经在很多领域得到了广泛应用。然而,粮情数据是十分复杂的,既有静态数据,又有动态数据,数据量庞大,数据的属性多样、特征多样,基于传统机器学习方法的预测模型很难对多源的大量粮情数据做出有效处理,从而导致这类预测模型通常只能采用部分数据对仓储粮情进行片面的预测,难以满足对“仓储粮情风险点”预测的实际需要。目前,深度学习算法得到了很大的发展和广泛的应用,其突出的优势在于可以自主地学习样本中隐含的深层次抽样特征,能够大大提高模型预测性能,从而成为构建仓储粮情风险点预测模型的一个不错的选择^[7-8]。

神经网络的发展得益于高性能图形处理单元(GPU)的应用,但 GPU 是较昂贵的计算资源,在面临大数据训练时,神经网络需要学习的权重参数非常多,且由于其较多的中间层,使得神经网络在应用 BP 算法进行训练时不仅十分耗时,而且也易于碰到梯度消失或梯度爆炸的问题。粮情数据属于大数据,数据不仅维数高、规模大,而且大多呈动态增长态势,这无疑给利用粮情大数据训练基于深度学习的粮情风险点预测模型带来了巨大的困难。因此,构建一种简洁的、易于训练的神经网络模型,使其能够在保证模型精度的前提下,适应大规模动态粮情数据的建模和训练是十分必要的。

宽度学习^[9-11]的提出从理论上为神经网络的出现提供了可能。采用宽度学习方式神经网络称为宽度学习系统(broad learning system, BLS),它是在原有浅层神经网络的基础上,把网络向宽度方向拓展,并且把隐藏层拓展为由特征节点和增强节点组成的级联结构。宽度学习系统的训练通过对隐藏层到输出层的权重求伪逆完成,大大快于梯度下降法的求解速度,同时,宽度学习系统还可采用增量学习的方法快速更新模型^[12]。由于宽度学习系统的这些优势,使其一经提出就受到重视,并已被成功应用于图像分

类^[13-14]、模式识别^[15]、电力预测^[16]、智能交通^[17]等领域。作者基于现有的宽度学习基本理论,采用基于宽度学习的特征提取与融合方法,以及基于增量学习的训练方法,结合粮情数据的多模态特征,提出了基于宽度学习系统的粮情风险预测模型,有效地改善了深度学习方法模型复杂、训练速度慢、所需数据集规模大等不足,可作为粮情风险预测的一个可行的替代方案。

1 宽度学习系统

1.1 宽度学习系统基本网络结构

宽度学习系统(BLS)是在 RVFLNN(随机向量函数链式神经网络)^[18]的基础上产生的,其网络结构如图 1 所示。BLS 的网络结构仍沿袭了传统的浅层神经网络的 3 层结构,不同的是 BLS 对其隐藏层进行了宽度方向的拓展,即隐藏层由多组特征节点和增强节点组成。需要说明的是,输入数据 X 不是直接与特征节点相连,而是经过某种映射后作为特征节点的值。

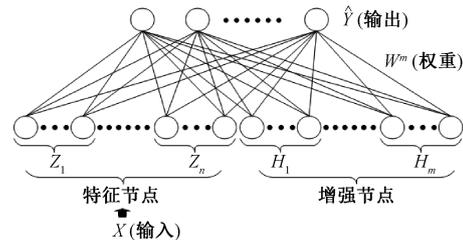


图 1 宽度学习系统的基本网络结构

Fig. 1 Basic structure of broad learning system

1.2 输入层

在 BLS 输入层,神经元个数等于输入数据的属性数。如输入数据有 M 个属性,则第 i 个输入数据可表示为向量 $x_i = (x_{i1}, x_{i2}, \dots, x_{iM})$ 。如果一共有 N 条这样的输入数据,则有 N 个 M 维的输入向量,可用矩阵表示为 $X = (x_1, x_2, \dots, x_i, \dots, x_N)^T$ 。BLS 对输入数据 X 进行批量处理,送入特征节点进行变换。由于输入数据的各属性具有不同的类型,为了方便处理,输入的原始数据一般都要经过归一化处理。

1.3 隐藏层

1.3.1 特征节点

由图 1 可知,BLS 的隐藏层是一个由特征节点层和增强节点层构成的级联结构。设其特征节点层由 n 组节点 Z_1, Z_2, \dots, Z_n 组成,且第 Z_i 组节点又由 q 个神经元组成,则输入数据 X 经过 Z_i

的特征映射后可表示:

$$Z_i = \varphi_i(XW_{ei} + \beta_{ei}), i = 1, 2, \dots, n, \quad (1)$$

式中: φ_i 是激活函数, 可以选用常见的 Sigmoid 函数、ReLU 函数等, 各组映射的激活函数可以不同; $W_{ei} \in R^{M \times q}$ 是网络的权重矩阵, 初始值可随机产生; $\beta_{ei} \in R^{N \times q}$ 是偏置矩阵, 初始值也是随机的, 并且这两个矩阵可通过稀疏自编码器微调以提取输入数据中更为稀疏的特征。

1.3.2 增强节点

BLS 的增强节点层可表示为 H_1, H_2, \dots, H_m , 假设其中的第 j 组增强节点 H_j 包含了 r 个神经元, 则有来自特征节点层的矩阵 Z^n , 经增强节点层后可得:

$$H_j = \xi_j(Z^n W_{hj} + \beta_{hj}), j = 1, 2, \dots, m, \quad (2)$$

式中: ξ_j 为激活函数, 可以选用 Sigmoid 函数、ReLU 函数等; $W_{hj} \in R^{N \times r}$ 为权重矩阵; $\beta_{hj} \in R^{N \times r}$ 为偏置矩阵, 其初始值均为随机值。

1.4 输出层

BLS 的输出层既可以处理回归问题也可以处理分类问题。对于分类问题, BLS 可采用独热码进行标签编码。如输入样本 x_i 的标签为 $y_i = (y_{i1}, y_{i2}, \dots, y_{iQ})$, Q 为类别数, 则独热码 $[0, 0, 1, \dots, 0]$ 表示输入样本是第 3 类。所有的标签可表示:

$$Y = (y_1, y_2, \dots, y_i, \dots, y_N)^T, \quad (3)$$

式中: Y 为标签矩阵。

输入样本 X 经过 BLS 后, 得到输出概率矩阵:

$$\hat{Y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_i, \dots, \hat{y}_N)^T, \quad (4)$$

式中: \hat{y}_i 表示样本 x_i 的标签是 y_i 的概率。

1.5 BLS 的建模和求解

在 BLS 中, 只有权重矩阵 W^m 需要训练, 减少了 BLS 的训练量。BLS 的训练基于以下目标函数:

$$\operatorname{argmin}_{W^m} \left(\|Y - \hat{Y}\|_2^2 + \frac{\lambda}{2} \|W^m\|_2^2 \right), \quad (5)$$

式中: $\|Y - \hat{Y}\|_2^2$ 表示训练误差, 训练过程是使其最小化的过程; λ 为正则化系数, 取值范围为 $0 \sim 1$; $\frac{\lambda}{2} \|W^m\|_2^2$ 是正则化项, 用于防止模型过拟合。

2 BLS 的改进——增量学习

增量学习是指当有新增样本时, 模型必须重新进行训练, 而只需针对新增数据做增量的训练

或局部的更新, 这种训练方式可以使模型的更新更快捷, 模型更新的代价也更小。文献[19]提出了动态更新的神经网络训练算法, 该算法通过求取伪拟矩阵^[20], 给出了在有新增输入数据时, 连接权重的快速更新方法。此后, 增量学习也应用到了随机森林^[21]、SVM^[22]等机器学习算法中, 逐渐成为机器学习领域中的一种重要的学习方法。将增强节点增量算法(EIBLS)和输入数据增量算法(IIBLS)用于改进传统宽度学习系统, 以提高 BLS 模型分类精度。

2.1 EIBLS 算法

EIBLS 算法是保持网络中特征节点数量不变, 增加增强节点的数量, 以达到提升模型特征提取能力的学习算法, 算法示意如图 2 所示。

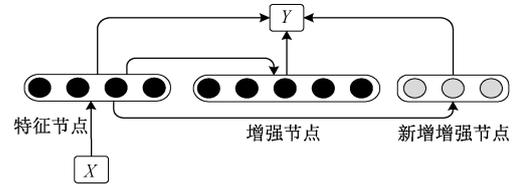


图 2 BLS 的增强节点增量算法

Fig. 2 Enhanced node incremental algorithm of BLS

当模型增加了若干个增强节点, 模型的隐藏层矩阵将由 $A = [Z^n | H^m]$ 扩展为 $A^{m+1} = [A | \xi(Z^n W_{m+1} + \beta_{m+1})]$, 其中 W_{m+1} 和 β_{m+1} 是新增的权重矩阵和偏置矩阵, ξ 是激活函数, 符号“|”表示两矩阵并置。由伪拟运算得到隐藏层矩阵的伪拟矩阵:

$$(A^{m+1})^+ = \begin{bmatrix} A^+ - DB^T \\ B^T \end{bmatrix}, \quad (6)$$

式中: $D = A^+ \xi(Z^n W_{h_{m+1}} + \beta_{h_{m+1}})$; $B^T = \begin{cases} C^+, & C \neq 0 \\ (1 + D^T D)^{-1} B^T A^+, & C = 0 \end{cases}$, $C = \xi(Z^n W_{h_{m+1}} + \beta_{h_{m+1}}) - AD$, C^+ 为矩阵 C 的伪逆。

根据式(5)、(6), 将隐藏层与输出层的连接权重 W^{m+1} 更新:

$$W^{m+1} = \begin{bmatrix} W^m - DB^T Y \\ B^T Y \end{bmatrix}. \quad (7)$$

2.2 IIBLS 算法

传统深度学习模型在输入数据增加时一般需要重新训练模型, 费时费力; 宽度学习模型在处理这一问题上表现了特殊的优势, 可以采用 IIBLS 算法快速更新模型。

如图 3 所示, 设 BLS 新增的数据为 X_n, A_n^m 为

原有的隐藏层矩阵,包括 n 组特征映射和 m 组增强映射。因新增的数据导致隐藏层更新:

$$A_x = [\varphi(X_a W_{e_1} + \beta_{e_1}), \dots, \varphi(X_a W_{e_n} + \beta_{e_n}) | \xi(Z_x^n W_{h_1} + \beta_{h_1}), \dots, \xi(Z_x^n W_{h_m} + \beta_{h_m})], \quad (8)$$

式中: $Z_x^n = [\varphi(X_a W_{e_1} + \beta_{e_1}), \dots, \varphi(X_a W_{e_n} + \beta_{e_n})]$ 是原有 BLS 特征节点的更新部分; $W_{e_i}, W_{h_j}, \beta_{e_i}, \beta_{h_j}$ 是随机生成的矩阵。更新后的隐藏层矩阵:

$${}^x A_n^m = \begin{bmatrix} A_n^m \\ A_x^T \end{bmatrix}, \quad (9)$$

对应的伪逆矩阵:

$$({}^x A_n^m)^+ = [(A_n^m)^+ - B D^T | B], \quad (10)$$

式中: $D^T = A_x^T (A_n^m)^+$;

$$B^T = \begin{cases} C^+, & C \neq 0 \\ (1 + D^T D)^{-1} (A_n^m)^+ D, & C = 0 \end{cases}$$

$$C = A_x^T - D^T A_n^m.$$

将最终权重 ${}^x W_n^m$ 更新:

$${}^x W_n^m = W_n^m + (Y_a^T - A_x^T W_n^m) B, \quad (11)$$

式中: Y_a^T 为新增数据 X_a 对应的标签。

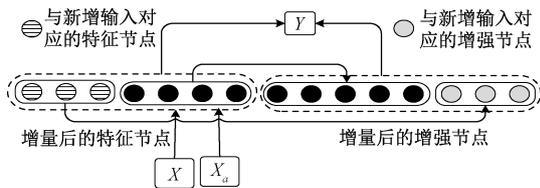


图 3 BLS 的输入数据增量算法

Fig. 3 Input data increment algorithm of BLS

3 基于 BLS 的粮情风险预测模型

基于 BLS 的粮情风险预测模型框架如图 4 所示。仓储粮情数据是以温度、湿度、气体组成、虫情等多种模态形式存在的。通过融合多种模态的信息进行机器学习是模型构建的必要步骤。但现有宽度学习系统及其改进算法在进行模式分类或预测时还都是只针对单模态的数据。结合粮情数据的多模态特征,在宽度学习系统现有框架的基础上,提出基于宽度学习系统的粮情风险预测模型,模型如图 4 所示。该模型旨在通过学习和融合多种模态的数据以提高预测精度。

由图 4 可知,粮情风险预测模型在宽度学习系统的基础上增加了典型相关性分析(CCA)模块和特征融合节点。首先对多种模态的粮情数据进行初步的特征提取,然后输入到相关性分析

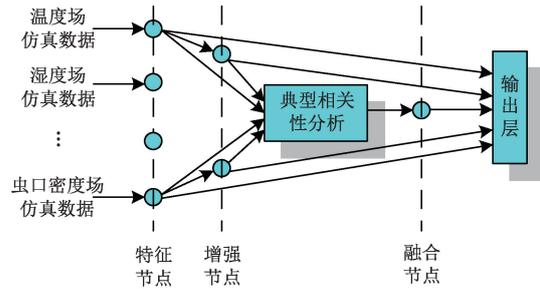


图 4 仓储粮情风险预测模型框架

Fig. 4 Framework of warehouse grain situation risk prediction model

模块进行相关性学习,再经过特征融合节点进行特征的融合,最后经输出层实现模式分类。该模型的优势在于相比于单一模态的分类模型,可以通过充分利用多模态的融合特征得到更高的分类性能。

CCA 可以对两个数据集进行融合降维,把两个不同模态的数据集按照某一关联规则映射到同一个数据空间,如果是多个模态的数据集,可以采用级联的方式进行融合和映射。CCA 的数学定义如下:

假设两个 n 维的数据集 $X = \{x_1, x_2, \dots, x_n\}$, $Y = \{y_1, y_2, \dots, y_n\}$,通过映射基向量 u 和 v 的变换后,分别得到新的坐标空间,定义数据集 X 和 Y 的相关性参数:

$$\rho = \max_{u,v} \frac{u^T \sum_{xy} v}{\sqrt{u^T \sum_{xx} u} \sqrt{v^T \sum_{yy} v}}, \quad (12)$$

式中: \sum_{xy} 和 \sum_{xx}, \sum_{yy} 分别为类间和类内协方差矩阵,即

$$\begin{cases} \sum_{xy} = E[xy^T] = \frac{1}{n} \sum_{i=1}^n x_i y_i^T \\ \sum_{xx} = E[xx^T] = \frac{1}{n} \sum_{i=1}^n x_i x_i^T \\ \sum_{yy} = E[yy^T] = \frac{1}{n} \sum_{i=1}^n y_i y_i^T \end{cases} \quad (13)$$

式(13)的相关计算实际上是一个优化问题,可以转化为特征值的求解问题。

假定模型的输入样本数为 N ,其特征节点和增强节点个数分别为 N_1 和 N_2 ,则一种模态数据的特征表达式: $X_1 = [Z_1 | H_1]$,它由一个宽度学习单元生成,其特征节点和增强节点分别可表示: $Z_1 = \{z_i | z_i \in R^{N_1}, i = 1, \dots, N\}$ 和 $H_1 = \{h_j | h_j \in R^{N_2}, j = 1, \dots, N\}$;同理,另一种模态数

据的特征表达式、特征节点和增强节点可以表示: $X_2 = [Z_2 | H_2]$, $Z_2 = \{z_i | z_i \in R^{N_2}, i = 1, \dots, N\}$ 和 $H_2 = \{h_j | h_j \in R^{N_2}, j = 1, \dots, N\}$, 由另一个宽度学习单元生成。为了更好地学习两种模态数据的共同特征, 需要先将它们混合再输入融合节点层。考虑到处理的方便性, 仅简单地将两种模态的数据并联起来, 作为最后提取的总特征和融合节点的净输入, 即

$$F^{N \times 2(N_1 + N_2)} = [X_1^{N \times (N_1 + N_2)} | X_2^{N \times (N_1 + N_2)}] \quad (14)$$

显然, F 就是融合节点层的新输入。为了更好地融合不同模态的数据, 引入融合节点映射层, 借鉴传统神经网络的非线性拟合将不同模态特征抽象融合, 最后利用输出矩阵进行特征学习, 以提高模型分类性能。假定融合层节点的个数为 N_3 , 则融合节点层输出:

$$T^{N \times N_3} = \phi(F^{N \times 2(N_1 + N_2)} \cdot W_i^{2(N_1 + N_2) \times N_3} + b_i^N \times N_3) \quad (15)$$

式中: $\phi(\cdot)$ 是一个 S 型非线性激活函数; W_i 为融合节点层的权重; b_i 表示融合节点的偏置(偏移量); t 对应前面 T 均表示输出。

4 试验结果与分析

4.1 数据获取与处理

试验数据来源于岳阳市某国家级储备库 2019—2020 年粮情测控数据。该粮库安装了全套的综合粮情测控系统, 可测储粮温度、湿度、虫情和仓内气体成分, 测控软件能以可视化的方式动态显示仓内温湿度云图、走势图等。试验共获取温湿度数据 735 条(每条数据包括仓内平均温度 1 项、粮堆各层平均温度、最高最低温度 12 项、仓内湿度 1 项、仓外温湿度各 1 项), 粮情报表 105 个, 温湿度云图图像 76 个, 虫情报表 20 个, 虫口密度图 24 个, 气体成分报告 10 个。部分数据样本如图 5 所示。

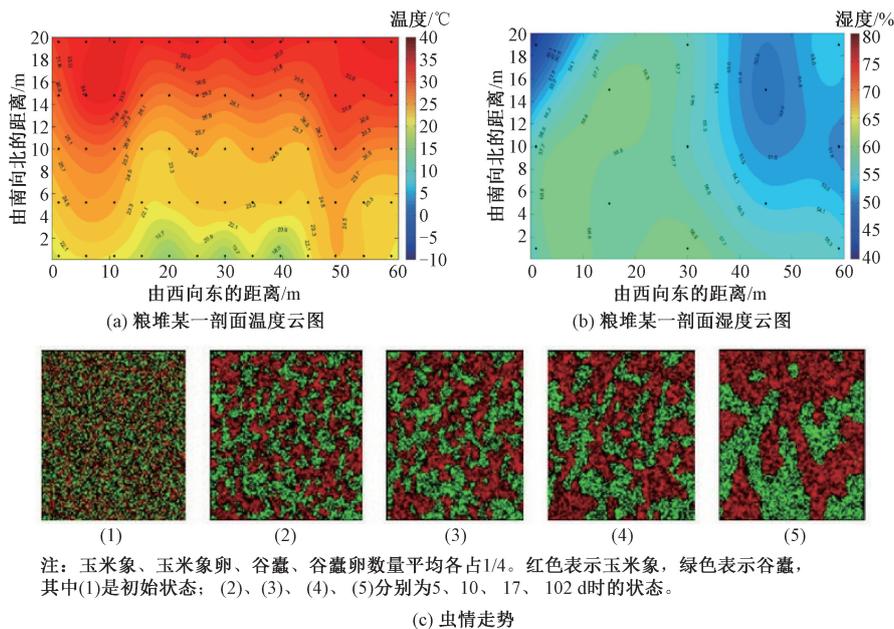


图 5 粮情多模态数据部分样本

Fig. 5 Partial sample of grain condition multi-modal data

显然, 获取的数据有两种类型(文本数据和图像数据)和多个模态(每个模态对应一种数据来源), 其中图像数据由文本数据可视化生成。对文本数据进行预处理, 将数据导出, 剔除病态数据, 形成温度、湿度、虫害、气体成分 4 种无格式纯文本文件, 共计 23 275 个数据, 将其中 70% 的数据作为训练集并标签化。图像数据 100 个, 采用 SamplePairing 扩增至 600 个, 采用 Labeling 工具加标签。本文研究目的在于预测粮

情风险等级, 在预测前, 将粮情风险划分为 4 个等级: 正常、低风险、中风险和高风险。各风险等级定义如表 1 所示。

需要说明的是, 利用表 1 给出的粮情风险等级有其局限性, 如夏季刚入仓的粮食温度可能高于正常指标。试验是通过人工的分析判断, 形成了数据集各样本标签。数据集中各样本是以天为单位, 1 d 的粮情数据为 1 个样本。

将数据分为 2 组, 文本数据和图像数据各一

表 1 粮情风险等级划分

Table 1 Risk grading of grain situation

等级名称	等级说明
正常	粮情各项指标处于正常范围,无须干预
低风险	粮情个别指标异常,但偏离不大,需要密切关注
中风险	粮情多项指标异常或个别指标偏离正常值较大,需要进一步监测评估,必要时进行干预
高风险	粮情多项指标偏离正常较大,粮情处于危险状态,需要立即干预

组,均由多种模态数据组成。将多模态的数据作为宽度学习系统网络模型的输入,通过图 4 所示的基于典型相关性分析模块,以提取到更深层次的粮情风险特征,从而提高模型的精准性。

由于粮情风险预测数据集是典型的不平衡数据集,正常数据占了大多数。为了提高以上各种风险数据占的比重,保证各类型数据的平衡,采用了 COMSOL Multiphysics 软件,对粮仓生态环境进行了仿真,并通过改变初始条件和人为设置参数,分别对低风险、中风险和高风险粮情进行了推演,取得了大量少数类型数据,并添加到训练数据集中,保证了数据集中数据类型的平衡性。最后,把数据按照 8:2 分成训练集和测试集。

4.2 评价指标

准确率作为分类器评价指标对于非平衡数据有失公平,为了客观评价分类算法性能好坏,根据混淆矩阵评估算法性能。根据表 2 的混淆矩阵,引入查全率 R (Recall)、查准率 P (Precision)、 F_1 -value 值 (F_1)、 G -mean 值等定义。

表 2 混淆矩阵

Table 2 Confusion matrix

类别	预测为正类	预测为负类
实际为正类	TP (True Positive)	FN (False Negative)
实际为负类	FP (False Positive)	TN (True Negative)

各指标定义如下:

$$R = \frac{TP}{TP + FN}, \quad P = \frac{TP}{TP + FP},$$

$$F_1 = \frac{2RP}{R + P},$$

$$G\text{-mean} = \sqrt{\frac{TP}{TP + FN} \times \frac{TN}{TN + FP}},$$

式中: F_1 是查全率与查准率加权调和平均值; G -mean 是正类准确率与负类准确率的综合指标。

4.3 试验设置

为了验证本文算法在小样本数据集上的性

能,试验的初始训练数据仅为 1 000 个,然后样本每次增加 1 000 个,评估宽度学习的增量学习算法对模型的准确率的影响。试验的目的在于检验本文的宽度学习模型在保证预测准确率的前提下是否可以作为深度学习的替代方案。因此,本文与一些深度学习的方法进行了对比,主要在识别率、训练和测试时间以及参数的敏感性等方面做了评估。

试验中对比了基本 BLS、级联特征映射的宽度学习 (CFBLS)^[13]、有限连接的级联特征映射宽度学习 (LCFBLS)^[13] 方法。CFBLS 改进了宽度学习模型的特征节点连接方式,采用了一种级联方式的连接结构。LCFBLS 与 CFBLS 结构相似,但最后一组特征节点用来生成增强节点。同时,本文的增量式宽度学习方法还与长短期记忆网络 (BiLSTM) 模型、Transformer 网络模型、SSD 模型和 YOLOv3 等 4 种深度学习方法进行了对比。宽度学习算法 BLS 的初始参数设置为 10×10 个特征节点和 100 个增强节点,增强节点更新 5 次,每次增加 100 个增强节点。

损失函数采用直接分类损失,该损失是约束模型训练的最基础的损失,通过交叉熵损失衡量由全连接层实现的分类器的输出,其计算方法见式 (16),其中, q_i 在 i 与真实标签 y 对应时取 1,否则取 0。

$$\text{Loss}_{\text{cross}} = \frac{1}{N} \sum_{i=1}^N -q_i \ln(p_i)。 \quad (16)$$

为了避免特征提取网络对当前任务产生过拟合而不适应新的检测任务或受到标记不准确的标签影响,在计算上述交叉熵损失时,引入了标签平滑策略,将待预测样本属于另一类的可能性也纳入衡量范围内,对交叉熵损失中的 q_i 进行修改。

$$q_i = \begin{cases} 1 - \varepsilon(N - 1)/N, & i = y \\ \varepsilon/N, & i \neq y \end{cases}。 \quad (17)$$

4.4 试验结果与分析

4.4.1 第 1 组数据试验结果与分析

为了证明提出的增量式宽度学习模型预测的有效性,利用文本数据对粮情风险预测模型进行建模,并比较采用 EIBLS 增量学习算法和 IIBLS 增量学习算法模型的预测精度。结果如表 3 和表 4 所示。

由表 3 和表 4 可知,随着在宽度学习中引入增量学习,模型的预测精度 (F_1) 逐渐提高。由于

表 3 EIBLS 增量学习算法的结果

Table 3 Results of EIBLS

特征节点	输入数据量	增强节点数	F_1	G -mean	训练时间/s	测试时间/s
10×10	1 000	100	0.83	0.88	8.84	0.12
10×10	1 000	200	0.84	0.89	9.23	0.12
10×10	1 000	300	0.85	0.90	11.34	0.13
10×10	1 000	400	0.85	0.90	13.78	0.13
10×10	1 000	500	0.87	0.91	15.72	0.14
10×10	1 000	600	0.88	0.91	16.48	0.15

表 4 IIBLS 增量学习算法的结果

Table 4 Results of IIBLS

特征节点	输入数据量	增强节点数	F_1	G -mean	训练时间/s	测试时间/s
10×10	1 000	100	0.83	0.88	8.84	0.12
10×10	2 000	200	0.88	0.93	10.46	0.12
10×10	3 000	300	0.91	0.94	22.51	0.13
10×10	4 000	400	0.91	0.94	31.49	0.13
10×10	5 000	500	0.94	0.96	45.43	0.14
10×10	6 000	600	0.97	0.98	48.51	0.15

IIBLS 增量学习算法能为预测模型提供新的信息,在增强节点相同的条件下,比仅增加增强节点的 EIBLS 算法具有更好的预测精度。如在增强节点都是 600 的条件下,输入数据为 6 000 时的 IIBLS 模型预测精度比输入数据为 1 000 时的 EIBLS 提高了 10%左右。但同时,输入数据的增加,带来了数据处理的开销增加,使得 IIBLS 模型的训练时间长于相同增强节点下的 EIBLS 的训练时间。

比较引入了增量学习的宽度学习模型与基本宽度学习模型、其他改进宽度学习模型和两种深度模型(BiLSTM 和 Transformer)在预测准确率、模型训练时间上的结果,BiLSTM 参数设置如下:学习率 0.001,卷积核窗口为(1,2,3,4),其中的每个 LSTM 大小为 128,采用 Adam 优化器;Transformer 的层数设为 4,自注意力头的个数为 8。使用相对位置编码长度为 8,batch size 数量为 32,训练期间使用初始学习率为 0.000 4 的 Ranger 优化器,每迭代 8 个 epoch 学习率衰减为原来的 4/5,训练过程中最大迭代次数设为 30。在解码过程中使用 Beam Search,集束大小设为 3。训练数据量从 1 000 增加到 6 000,对比结果如表 5 所示。

由表 5 可知,综合了预测的精确率和训练时间,本文提出的方法,优于基本宽度学习模型和两种改进宽度学习模型,特别是本文模型在训练时间上,由于采用了增量学习的算法,训练时间大大减少。与两种常见的深度学习算法进行比较,IIBLS 的预测精度和对比的两种深度学习方

表 5 本文模型与对比模型在第 1 组数据上的预测效果对比

Table 5 Comparison between the proposed model and the other models on the first dataset

预测模型算法	F_1	G -mean	训练时间/s
BiLSTM	0.87	0.89	722.67
Transformer	0.89	0.90	605.22
BLS	0.82	0.84	17.78
CFBLS	0.84	0.86	19.56
LCFBLs	0.86	0.88	20.33
EIBLS	0.87	0.88	16.48
IIBLS	0.88	0.89	48.51

法非常接近,EIBLS 方法也只有很小的差距,但本文的两种宽度学习模型训练时间大大少于两种深度学习模型。由此可以看出,本文的宽度学习模型在粮情风险预测上完全可以作为深度模型的替代模型。

4.4.2 第 2 组数据试验结果与分析

利用图像数据对粮情风险预测模型进行建模,并比较采用 EIBLS 增量学习算法和 IIBLS 增量学习算法模型的预测精度。所有图像被调整为 224×224 的分辨率,宽度学习网络基本模型感知域为 4×4,特征图和池大小均为 3。对比结果如表 6 和表 7 所示。

由表 6 和表 7 可知,对于图像数据,随着增量学习的引入,模型的预测精度提高得更快,且在增强节点相同的条件下,IIBLS 增量学习算法比 EIBLS 增量学习算法有更高的预测准确率。但由

表 6 EIBLS 增量学习算法的结果

Table 6 Results of EIBLS

特征节点	输入数据量	增强节点数	F_1	G -mean	训练时间/s	测试时间/s
10×10	100	100	0.78	0.81	27.53	0.33
10×10	100	200	0.79	0.82	29.38	0.33
10×10	100	300	0.81	0.83	32.74	0.34
10×10	100	400	0.81	0.83	34.67	0.35
10×10	100	500	0.83	0.85	35.14	0.35
10×10	100	600	0.84	0.86	37.65	0.36

表 7 IIBLS 增量学习算法的结果

Table 7 Results of IIBLS

特征节点	输入数据量	增强节点数	F_1	G -mean	训练时间/s	测试时间/s
10×10	100	100	0.78	0.81	27.53	0.33
10×10	200	200	0.83	0.85	34.55	0.33
10×10	300	300	0.88	0.92	52.78	0.34
10×10	400	400	0.91	0.95	63.22	0.34
10×10	500	500	0.94	0.96	70.67	0.36
10×10	600	600	0.96	0.98	79.15	0.36

于图像处理比较费时,模型的训练时间稍长于文本输入数据下的训练时间。

为了证明增量学习式 BLS 预测模型对图像数据的有效性,将 EIBLS 和 IIBLS 两种预测模型与基本 BLS、CFBLS、LCFBLs 以及两种深度神经网络图像检测模型 SSD、YOLOv3 在预测性能上进行对比,结果如表 8 所示。各对比模型参数设置如下:IIBLS 的初始结构设置为 4×4 感知域、100 个增强节点和 10×10 特征节点、特征映射和池大小为 3;将 BLS、CFBLS、LCFBLs 的初始结构设置为 1 000 个增强节点和 100×10 特征节点。

表 8 本文模型与对比模型在第 2 组数据上的预测效果对比

Table 8 Comparison between the proposed model and the other models on the second dataset

预测模型算法	F_1	G -mean	训练时间/s
YOLOv3	0.92	0.95	4 647.13
SSD	0.91	0.95	5 427.82
BLS	0.78	0.84	29.43
CFBLS	0.79	0.86	30.21
LCFBLs	0.81	0.88	32.56
EIBLS	0.88	0.89	37.65
IIBLS	0.90	0.93	79.15

由表 8 可知,本文提出的 BLS,对于图像数据,在预测的精确率和训练时间上,整体都优于基本宽度学习模型和两种改进宽度学习模型。与两种常见的一阶段目标检测深度学习网络模型相比,IIBLS 和 EIBLS 方法只有很小的差距,但训练时间大大少于两种典型深度学习模型。由

此可以看出,本文的宽度学习模型在粮情风险预测上,无论是针对文本数据还是图像数据,完全可以作为深度模型的替代。

5 结论

基于近年来出现的宽度学习基本理论,通过引入增量学习方法和相关性分析方法,建立了一个基于改进宽度学习网络的仓储粮情风险点预测模型。该模型针对深度学习模型构建时存在的训练时间长、资源需求大等问题,在保证预测精度不降低的前提下,大大降低了模型的训练难度和资源消耗,使得精准的仓储粮情风险预测模型在资源配置较低的储粮企业也可以得到实现,可为企业减少大量的信息化投资,降低储粮的成本。目前,本文的宽度学习模型在预测精度上作为一种深度学习的替代方法,仍需要进一步提高预测精度。下一步的工作将集中在如何提高不同模态粮情数据的融合效果,以及如何在宽度学习中引入注意力机制,以进一步提高宽度学习模型的预测精度和应用范围。

参考文献:

- [1] 苑江浩,郑丹,孟祥雪,等. 粮情监测预警技术研究进展[J]. 中国粮油学报, 2022, 37(11):19-26.
- [2] 麦嘉琪,徐鹏程,丁松,等. 支持向量回归

- 预测不锈钢的点蚀电位 [J]. 上海大学学报(自然科学版), 2022, 28(3): 485-491.
- [3] 王烽, 李泽平, 林川, 等. 基于随机森林的带宽预测算法研究与实现 [J]. 计算机工程与设计, 2020, 41(7): 1892-1898.
- [4] 闫祥祥. 使用 ARIMA 模型预测公园绿地面积 [J]. 计算机科学, 2020, 47(S2): 531-534, 556.
- [5] 苏琪, 王海波, 施晓辰, 等. 基于灰色预测模型的参数寻优方法及能源预测应用 [J]. 南昌大学学报(理科版), 2022, 46(3): 371-378.
- [6] 赵锐, 赵丽萍, 陈静芳, 等. 基于 BP 神经网络的垃圾渗滤液输运管道结垢趋势预测 [J]. 中国测试, 2022, 48(7): 1-7.
- [7] 赵岩, 周俊冬, 任双双. 基于深度学习算法的粮堆温度预测研究 [J]. 粮食科技与经济, 2019, 44(11): 52-56.
- [8] 段珊珊, 杨卫东, 肖乐, 等. 一种基于气象数据的仓储粮堆表层温度预测方法 [J]. 中国粮油学报, 2020, 35(2): 152-158.
- [9] PHILIP CHEN C L, LIU Z L. Broad learning system: an effective and efficient incremental learning system without the need for deep architecture [J]. IEEE transactions on neural networks and learning systems, 2018, 29(1): 10-24.
- [10] PHILIP CHEN C L, LIU Z L. Broad learning system: a new learning paradigm and system without going deep [C]//2017 32nd Youth Academic Annual Conference of Chinese Association of Automation (YAC). May 19-21, 2017, Hefei, China. IEEE, 2017: 1271-1276.
- [11] PHILIP CHEN C L, LIU Z L, FENG S. Universal approximation capability of broad learning system and its structural variations [J]. IEEE transactions on neural networks and learning systems, 2019, 30(4): 1191-1204.
- [12] 任长城, 袁超, 孙彦丽, 等. 宽度学习系统研究进展 [J]. 计算机应用研究, 2021, 38(8): 2258-2267.
- [13] 王召新, 续欣莹, 刘华平, 等. 基于级联宽度学习的多模态材质识别 [J]. 智能系统学报, 2020, 15(4): 787-794.
- [14] 叶荣珂, 孔庆辰, 李道亮, 等. 宽度学习的虾新鲜度检测方法 [J]. 光谱学与光谱分析, 2022, 42(1): 164-169.
- [15] 贾晨, 刘华平, 续欣莹, 等. 基于宽度学习方法的多模态信息融合 [J]. 智能系统学报, 2019, 14(1): 150-157.
- [16] 周楠, 徐潇源, 严正, 等. 基于宽度学习系统的光伏发电功率超短期预测 [J]. 电力系统自动化, 2021, 45(1): 55-64.
- [17] 朱玉斌, 延向军, 申旭奇, 等. 基于级联宽度学习的疲劳驾驶检测 [J]. 计算机工程与设计, 2020, 41(2): 537-541.
- [18] RAIGURU P, SAHANI M, ROUT S K, et al. A digital direction of arrival estimator based on fast on-line sequential random vector functional link network [J]. AEU-international journal of electronics and communications, 2021, 142: 153986.
- [19] CHEN C L P, WAN J Z. A rapid learning and dynamic stepwise updating algorithm for flat neural networks and the application to time-series prediction [J]. IEEE transactions on systems, man, and cybernetics, Part B (Cybernetics), 1999, 29(1): 62-72.
- [20] IPLEA F L, DRĂGOI V F. Generalized inverse based decoding [C]//2022 IEEE International Symposium on Information Theory (ISIT), Espoo, Finland, 2022: 2791-2796.
- [21] 王爱平, 万国伟, 程志全, 等. 支持在线学习的增量式极端随机森林分类器 [J]. 软件学报, 2011, 22(9): 2059-2074.
- [22] 王洪波, 赵光宙, 齐冬莲, 等. 一类支持向量机的快速增量学习方法 [J]. 浙江大学学报(工学版), 2012, 46(7): 1327-1332.

(责任编辑 周凤航)