基于长短期记忆生成对抗网络的小麦品质多指标预测模型

蒋华伟* 张 磊

(河南工业大学信息科学与工程学院 郑州 450001)

摘 要：小麦多生理生化指标变化趋势反映了储藏品质的劣变状态，预测多指标时序数据会因关联性及相互作用而产生较大误差，为此该文基于长短期记忆网络(LSTM)和生成式对抗网络(GAN)提出一种改进拓扑结构的长短期记忆生成对抗网络(LSTM-GAN)模型。首先，由LSTM预测多指标不同时序数据的劣变趋势；其次，根据多指标的关联性并结合GAN的对抗学习方法来降低综合预测误差；最后通过优化目标函数及训练模型得出多指标预测结果。经实验分析发现：小麦多指标的长短期时序数据的变化趋势不同，进一步优化模型结构及训练时序长度可有效降低预测结果的误差；特定条件下小麦品质过快劣变会使多指标预测误差增大，因此应充分考虑储藏期环境变化对多指标数据的影响；LSTM-GAN模型的综合误差相对于仅使用LSTM预测降低了9.745%，并低于多种对比模型，这有助于提高小麦品质多指标预测及分析的准确性。

关键词：长短期记忆网络；生成式对抗网络；小麦多指标；预测模型

中图分类号：TP391 **文献标识码**：A **文章编号**：1009-5896(2020)12-2865-08

DOI: 10.11999/JEIT190802

Multi-index Prediction Model of Wheat Quality Based on Long Short-Term Memory and Generative Adversarial Network

JIANG Huawei ZHANG Lei

(*College of Information Science and Engineering*, *Henan University of Technology*, *Zhengzhou* 450001, *China*)

Abstract: The change trend of multi-index of wheat reflects the deterioration state of storage quality, while the predicted multi-index data will produce large errors due to its correlation and interaction. For this reason, an improved Long Short-Term Memory and Generative Adversarial Network(LSTM-GAN) model is proposed. The deterioration trend of different time series data of multi-index is predicted by Long Short-Term Memory(LSTM) network, and the improved model may reduce comprehensive prediction error by using Generative Adversarial Network(GAN) according to the correlation of multi-index. Finally, the prediction results obtained by optimizing the objective function and model structure. The experimental analysis shows that the training sequence length and structural parameters of the optimization model can effectively reduce the error of the prediction result. The deterioration of wheat quality under certain conditions will increase the prediction error of multi-index. Therefore, the influence of environmental changes during storage period on multi-index data should be fully considered. The comprehensive error of the LSTM-GAN model is reduced by 9.745% compared with the LSTM prediction and lower than multiple comparison models, which can improve the prediction of wheat quality indexes.

Key words: Long Short-Term Memory(LSTM) network; Generative Adversarial Network(GAN); Wheat multi-index; Prediction model

1 引言

小麦籽粒品质随储藏时间延长逐渐产生劣变，同时反映在多指标不同程度的数值变化上，因此可通过研究小麦多指标时序数据的变化趋势来了解其储藏品质的劣变程度，并以此调控小麦的储藏过程。由于小麦储藏环境的差异性以及多指标间的不同相互作用[1,2]，多指标时间序列数据的预测存在一定的误差，且随着储藏时间延长其误差不断增大，进而影响到小麦品质评价的准确性。为此，需要探索更为有效的小麦品质多指标预测算法，以提高预测结果的准确性及稳定性，为小麦品质评价提供一定的技术支持。

收稿日期：2019-10-16；改回日期：2020-10-18；网络出版：2020-10-26
*通信作者： 蒋华伟 lhwcad@126.com
基金项目：国家自然科学基金(51677055)，河南省自然科学基金(162300410055)，河南省高校科技创新团队计划项目(16IRTSTHN026)
Foundation Items: The National Natural Science Foundation of China (51677055), The Natural Science Foundation of Henan Province (162300410055), The Science and Technology Innovation Team Planning Project of University of Henan Province (16IRTSTHN026)

近年来人工智能算法及神经网络模型迅速发展，有效地提高了数据挖掘及预测分析的效率和准确性，其中循环神经网络[3](Recurrent Neural Network, RNN)可以处理任意长度的时序数据，它的拓扑结构包含自反馈机制，且具有一定的记忆能力，适用于小麦多指标时序数据的预测分析。然而RNN在对长序列进行训练时易出现梯度消失或梯度爆炸，为解决以上问题，长短期记忆网络[4](Long Short-Term Memory, LSTM)通过引入门控单元来控制记忆的迭代速度，进一步提高了预测模型的效率及稳定性。后来Mahasseni等人[5]在实验中发现LSTM预测仍存在一定的指数级错误积累，为此，Yang等人[6]通过在LSTM单元输入层和激活函数间加入连接层，提出了对抗LSTM用以改善序列预测结果，为后续研究提供了理论基础。

以上方法虽然有助于提高数据预测结果的准确性，但是在小麦多指标数据预测中，多指标间的关联性及储藏环境的差异性仍会对预测结果产生一定的影响。本文提出一种改进的长短期记忆生成对抗网络(Long Short-Term Memory and Generative Adversarial Network, LSTM-GAN)模型，在充分考虑多指标时序变化特征的基础上，由LSTM对多指标的长短期记忆时序数据进行预测分析，并加入了多指标的对抗学习。利用生成式对抗网络(Generative Adversarial Network, GAN)中生成器和判别器博弈的方法进一步降低训练过程中多指标的综合预测误差，通过时序分析和模型结构优化，来提高小麦品质指标预测的准确性。

2 LSTM和GAN的模型原理

2.1 长短期记忆网络

记忆单元是LSTM的核心组件，每个LSTM单元包含一个元组，在t时刻其状态为$\boldsymbol{c}_t$，包含了序列的长期记忆信息；在t时刻隐含层的状态为$\boldsymbol{h}_t$，包含了序列的短期记忆信息。记忆单元信息的读取和修改是通过控制遗忘门、输入门以及输出门来实现的，如图1所示为LSTM单元的基本结构。

在时刻t时，LSTM记忆单元的输入包括：序列输入$\boldsymbol{x}_t$、记忆单元t–1时刻的状态$\boldsymbol{c}_{t-1}$以及隐含层t–1时刻的状态$\boldsymbol{h}_{t-1}$；其输出包括：t时刻记忆单元的状态$\boldsymbol{c}_t$与隐含层的状态$\boldsymbol{h}_t$。在t时刻假定输入层数据为$\boldsymbol{x}_t$，则此刻隐含层状态$\boldsymbol{h}_t$和输出层的预测值为$\boldsymbol{y}_t$为

$$\boldsymbol{h}_t = f\left(\boldsymbol{U}\boldsymbol{x}_t + \boldsymbol{W}\boldsymbol{h}_{t-1} + \boldsymbol{b}_h\right) \tag{1}$$

$$\boldsymbol{y}_t = g\left(\boldsymbol{V}\boldsymbol{h}_t + \boldsymbol{b}_y\right) \tag{2}$$

式中，$\boldsymbol{h}_{t-1}$为t–1时刻的隐含层状态；f和g分别表示

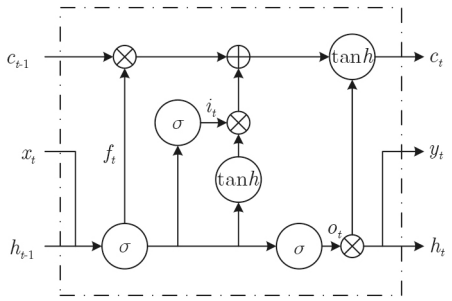

图 1 长短期记忆网络单元结构

隐含层和输出层的激活函数；$\boldsymbol{U}$和$\boldsymbol{V}$分别表示输入层与隐含层、隐含层与输出层之间的权重矩阵；$\boldsymbol{W}$表示隐含层中的自反馈权重矩阵；$\boldsymbol{b}_h$和$\boldsymbol{b}_y$分别表示隐含层和输出层的偏置项。则LSTM中遗忘门、输入门、输出门在t时刻的状态$\boldsymbol{f}_t$, i_t, o_t和记忆单元的状态$\boldsymbol{c}_t$与隐含层状态$\boldsymbol{h}_t$计算公式为

$$i_t = \sigma\left(\boldsymbol{W}_{xi}\boldsymbol{x}_t + \boldsymbol{W}_{hi}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{ci}\boldsymbol{c}_{t-1} + \boldsymbol{b}_i\right) \tag{3}$$

$$f_t = \sigma\left(\boldsymbol{W}_{xf}\boldsymbol{x}_t + \boldsymbol{W}_{hf}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{cf}\boldsymbol{c}_{t-1} + \boldsymbol{b}_f\right) \tag{4}$$

$$o_t = \sigma\left(\boldsymbol{W}_{xo}\boldsymbol{x}_t + \boldsymbol{W}_{ho}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{co}\boldsymbol{c}_t + \boldsymbol{b}_o\right) \tag{5}$$

$$\boldsymbol{c}_t = f_t\boldsymbol{c}_{t-1} + l_t\tanh\left(\boldsymbol{W}_{xc}\boldsymbol{x}_t + \boldsymbol{W}_{hc}\boldsymbol{h}_{t-1} + \boldsymbol{b}_c\right) \tag{6}$$

$$\boldsymbol{h}_t = o_t\tanh\left(\boldsymbol{c}_t\right) \tag{7}$$

式中，$\boldsymbol{W}_{xc}$, $\boldsymbol{W}_{xi}$, $\boldsymbol{W}_{xf}$和$\boldsymbol{W}_{xo}$为连接输入信号$\boldsymbol{x}_t$的权重矩阵；$\boldsymbol{W}_{hc}$, $\boldsymbol{W}_{ih}$, $\boldsymbol{W}_{hf}$和$\boldsymbol{W}_{ho}$为连接隐含层输出信号$\boldsymbol{h}_t$的权重矩阵；$\boldsymbol{W}_{ci}$, $\boldsymbol{W}_{cf}$和$\boldsymbol{W}_{co}$为连接神经元激活函数输出矢量$\boldsymbol{c}_t$和门函数的对角矩阵；$\boldsymbol{b}_i$, $\boldsymbol{b}_c$, $\boldsymbol{b}_f$和$\boldsymbol{b}_o$为偏置向量；tanh为双曲正切激活函数，σ表示sigmoid激活函数，其作用是将变量映射至区间[0,1]中。

在研究小麦产后品质变化规律时，传统LSTM网络主要采用已知的时序信息去预测指定时刻的数值，然而，这些小麦指标序列的计量单位、数值范围存在较大差异，单个生理生化指标在分析整体品质时并不具有代表性，因此，本文拟采用权值矩阵方法整合多个生理生化指标序列的LSTM预测结果，再采用组合优化算法进一步降低小麦多指标预测的综合误差。

2.2 生成式对抗网络

GAN是由Goodfellow等人[7]在2014年依据二元零和博弈提出的一种生成式模型，其框架中包含1组相互对抗的生成器和判别器模块，用以判断和监视模型学习效果。其中判别器是一个二分类模型，可用交叉熵计算目标函数

$$J\left(D\right) = -\frac{1}{2}\mathrm{E}_{x\sim p_{\mathrm{data}}(x)}\left[\ln D\left(x\right)\right] - \frac{1}{2}\mathrm{E}_{z\sim p_z(\boldsymbol{z})} \cdot \left[\ln\left(1 - D\left(G\left(\boldsymbol{z}\right)\right)\right)\right] \tag{8}$$

G和D分别表示生成器和判别器的可微函数，E是

目标函数的期望值，x是真实数据样本，$\boldsymbol{z}$是随机噪声矢量，$G(\boldsymbol{z})$是判别器的生成数据。第1项表示D判断出x是真实数据的情况，第2项则表示D判别出数据是由生成器G将噪声矢量$\boldsymbol{z}$映射而成的生成数据，G与D进行二元零和博弈，生成器G的目标函数$J(G)=-J(D)$。因此，GAN的优化问题可转化为极大极小博弈问题

$$\min_G \max_D V(D,G)=\mathrm{E}_{x\sim p_{\mathrm{data}}(x)}\left[\ln D\left(x\right)\right]+\mathrm{E}_{z\sim p_z(\boldsymbol{z})}\cdot\left[\ln\left(1-D\left(G\left(\boldsymbol{z}\right)\right)\right)\right] \tag{9}$$

因为G在训练初期所生成的数据不够逼真，所以D很容易将生成数据与真实数据区分开来，导致G误差梯度下降较为缓慢，因此通过最大化$\ln D(G(\boldsymbol{z}))$而非最小化$\ln(1-D(G(\boldsymbol{z})))$来训练$G$是一个更好的策略。生成器$G$采用神经网络训练其目标函数，而不是直接取$J(D)$的相反数，即使判别器$D$准确地拒绝了所有生成样本，$G$仍可以继续学习直至达到模型要求的效果，为此可以将极大极小博弈改为非饱和博弈

$$J\left(D\right)=-\frac{1}{2}\mathrm{E}_{x\sim p_{\mathrm{data}}(x)}\left[\ln D\left(x\right)\right]-\frac{1}{2}\mathrm{E}_{z\sim p_z(\boldsymbol{z})}\cdot\left[\ln\left(1-D\left(G\left(\boldsymbol{z}\right)\right)\right)\right] \tag{10}$$

$$J\left(G\right)=-\frac{1}{2}\mathrm{E}_{z\sim p_z(\boldsymbol{z})}\ln D\left(G\left(\boldsymbol{z}\right)\right) \tag{11}$$

GAN为对抗学习提供了比较有效的方法，已广泛应用于图像处理[8]等领域问题的解决。小麦多指标数据间存在着紧密的关联性，为此可根据GAN来权衡多指标时序数据在小麦整体品质分析中的不同表现作用，为小麦多指标数据预测及品质评价提供一定程度的数据支持。

3　LSTM-GAN小麦多指标预测模型

3.1　LSTM-GAN模型网络结构

由上述可知，尽管LSTM记忆单元可通过控制长短期记忆信息的更新来预测单个指标的时序变化规律，但是不能综合多个指标的数值变化来对小麦整体品质劣变趋势进行把握；GAN虽然可用于多指标的对抗学习以获得小麦品质的整体变化特征，然而GAN中生成器和判别器的网络结构如果选择不当会直接导致模型性能的下降。因此，本文改进LSTM和GAN网络来融合多指标整体特征的变化趋势，利用GAN中对抗学习的方法来检验小麦多指标数据集预测结果的误差大小，并找出误差较大部分进一步优化，由此提出一种LSTM-GAN模型来提高多指标预测分析的准确性。LSTM-GAN模型主要由生成器G、LSTM模块、判别器D组成，其网络拓扑结构如图2所示。

由图2可知，输入层将训练集、测试集的小麦多指标数据标准化后传给生成器G，其中小麦n个生理生化指标在m个测试间隔点测试获得的一系列数据作为n个时间序列，生成器G将这些时间序列数据加权组合初始化为一个$n\times m$的小麦特征信息矩阵，并分别传输给LSTM模块、判别器D进行后续计算；LSTM模块收到生成器G传来的小麦特征信息矩阵后将其拆分为n个时间序列，初始化神经网络中各记忆单元及隐含层的权值等，分别进行训练计算获得预测结果，并将这n个序列的预测结果组合为一个新的预测结果矩阵，用于判别器D计算预测综合误差；判别器D接收并存储生成器G传来的特征信息矩阵和LSTM模块传输的预测结果矩阵，计算出两个矩阵的概率分布差异作为该部分的目标函数。

如果判别器D计算得出两个矩阵中某些行列的差异较大，即本轮训练过程中一些小麦品质指标在相应储藏条件下的预测误差相对较大，则反馈给LSTM模块并调整优化相应的训练过程，使其得出更准确的预测结果后再次更新传给判别器D的预测结果矩阵。LSTM-GAN模型在判别器D不断纠正预测结果矩阵中相对误差较大的数据之后，整体损失函数也将会逐渐降低，直到模型预测结果的综合

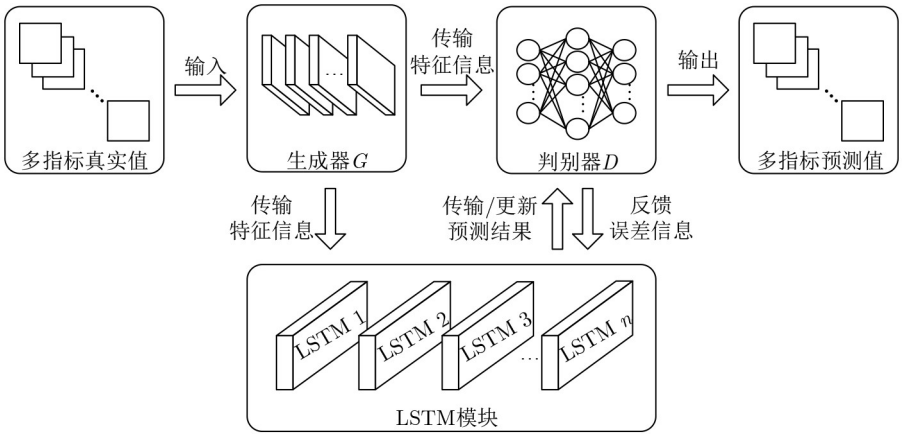

图 2 长短期记忆生成对抗网络

误差降低到一定范围内或者达到设置的训练次数时，终止模型训练过程并输出小麦多指标的预测值。

3.2 LSTM-GAN模型目标函数

由LSTM-GAN模型结构可知，生成器G的输入为小麦多生理生化指标数据集，其中小麦n个生理生化指标在m个测试间隔点的数据可直接构成一个$n\times m$的数据集矩阵，生成器G通过为这个数据集矩阵配置相应的权重进而可获得相应的小麦特征信息矩阵，由此得到的小麦特征信息矩阵$\boldsymbol{G}(t)$包含了数据集中多个指标的时序信息，因此，在t时刻生成器G输出的小麦特征信息矩阵$\boldsymbol{G}(t)$为

$$\boldsymbol{G}(t)=\sum_{i=1}^{n}\boldsymbol{W}_{it}\cdot x_t\sim p_t(\text{data}) \tag{12}$$

式中，$\boldsymbol{W}_{it}$为生成器G根据熵权法[9]通过衡量各指标序列的互信息量所得到的对应权值，x_t为小麦多指标在t时刻的采样数据，它服从于样本中真实指标数据的分布$p_t(\text{data})$。另外，在训练过程中，若小麦特征信息矩阵与多指标真实数据分布之间存在较大差异，可使用熵权法对$\boldsymbol{W}_{it}$进行更新，以此来提高模型的训练速度，从而使得生成器G能够快速拟合小麦指标数据集的真实特征分布。

由式(8)可知，传统的GAN网络采用二元分类器判别生成数据是否来源于真实数据。为了获得适用于小麦多指标预测的对抗学习模型，本文改进GAN的模型结构，利用特征信息矩阵和预测结果矩阵的概率分布差异作为判别器D的判别目标，使其找到预测结果中误差相对较大的部分并进一步优化，其判别目标及优化函数为

$$\begin{aligned}J(D)=&-\frac{1}{2}\mathrm{E}_{t\sim g(\text{data})}\left[\ln D\left(\boldsymbol{G}(t)\right)\right]-\frac{1}{2}\mathrm{E}_{t\sim l(\text{data})}\\&\cdot\left[\ln\left(1-D\left(\boldsymbol{L}\left(\boldsymbol{G}(t)\right)\right)\right)\right]\end{aligned} \tag{13}$$

$$\begin{aligned}\min_G\max_D V(G,D)=&\mathrm{E}_{t\sim g(\text{data})}\left[\ln D\left(\boldsymbol{G}(t)\right)\right]+\mathrm{E}_{t\sim l(\text{data})}\\&\cdot\left[\ln\left(1-D\left(\boldsymbol{L}\left(\boldsymbol{G}(t)\right)\right)\right)\right]\end{aligned} \tag{14}$$

对于生成器G输出的小麦特征信息矩阵，由LSTM模块划分为多个数据序列，通过控制有效记忆信息的取舍以及减小无关信息所产生的影响，以提高这些序列预测结果的效率和准确度，由LSTM模块训练得到的预测结果矩阵为$\boldsymbol{L}(\boldsymbol{G}(t))$。LSTM模块单元的优化目标是取得更小的误差，由式(14)可以得出本文LSTM-GAN模型计算的整体目标函数为

$$\begin{aligned}&\min_{G,L}\max_D V(G,L,D)\\&=\mathrm{E}_{t\sim g(\text{data})}\left[\ln D\left(\boldsymbol{G}(t)\right)\right]\\&\quad+\mathrm{E}_{t\sim l(\text{data})}\left[\ln\left(1-D\left(\boldsymbol{L}\left(\boldsymbol{G}(t)\right)\right)\right)\right]\\&=\int\left(x\sim p_{\text{data}}\boldsymbol{G}(t)\ln(D(\boldsymbol{G}(t)))\right.\\&\quad\left.+x\sim p_{\text{data}}(\boldsymbol{L}(\boldsymbol{G}(t)))\ln(1-D(\boldsymbol{G}(t)))\right)\mathrm{d}t\end{aligned} \tag{15}$$

由式(15)可知，LSTM-GAN模型的目标函数是在式(14)的基础上，把最小化LSTM网络预测误差作为训练目标。LSTM-GAN模型中生成器G、判别器D的优化目标与传统GAN网络一致，本文保留GAN的对抗学习并在此基础上分部优化LSTM网络以提高时序数据预测的准确性。由于对于任意的$(a,b)\in\mathrm{R}^2$且不等于0，函数$a\ln(y)+b\ln(1-y)$关于y的最大值为$a/(a+b)$，因此判别器D达到最优结果的输出为

$$D^*_{(G,L)}=\frac{x\sim p_{\text{data}}\boldsymbol{G}(t)}{x\sim p_{\text{data}}\boldsymbol{G}(t)+x\sim p_{\text{data}}\boldsymbol{L}(\boldsymbol{G}(t))} \tag{16}$$

本文LSTM-GAN模型中判别器D采用JS散度计算特征信息矩阵和预测结果矩阵之间的相似程度，以此来衡量两个概率分布之间的差异性。在给定最优判别器$D^*_{(G,L)}$的条件下，将式(16)代入式(15)，可将目标函数转换为最小化多指标的特征信息矩阵分布$P_{g(t)}$与预测结果矩阵分布$P_{l(g(t))}$之间的JS散度

$$\begin{aligned}&C(G,L)\\&=\max_D V(G,L,D)\\&=\mathrm{E}_{t\sim p_g}\ln[D^*_{(G,L)}\boldsymbol{G}(t)]\\&\quad+\mathrm{E}_{t\sim p_{l(g)}}\ln[1-D^*_{(G,L)}(\boldsymbol{L}(\boldsymbol{G}(t)))]\mathrm{E}_{t\sim p_g}\\&\quad\cdot\ln\frac{x\sim p_{\text{data}}\boldsymbol{G}(t)}{\frac{1}{2}[x\sim p_{\text{data}}\boldsymbol{G}(t)+x\sim p_{\text{data}}\boldsymbol{L}(\boldsymbol{G}(t))]}\\&\quad+\mathrm{E}_{t\sim p_{l(g)}}\ln\frac{x\sim p_{\text{data}}\boldsymbol{L}(\boldsymbol{G}(t))}{\frac{1}{2}[x\sim p_{\text{data}}\boldsymbol{G}(t)+x\sim p_{\text{data}}\boldsymbol{L}(\boldsymbol{G}(t))]}\\&\quad-2\ln 2\\&=\mathrm{KL}\left(x\sim p_{\text{data}}\boldsymbol{G}(t)\middle\|\right.\\&\quad\left.\frac{x\sim p_{\text{data}}\boldsymbol{G}(t)+x\sim p_{\text{data}}\boldsymbol{L}(\boldsymbol{G}(t))}{2}\right)\\&\quad+\mathrm{KL}\left(x\sim p_{\text{data}}\boldsymbol{G}(t)\middle\|\right.\\&\quad\left.\frac{x\sim p_{\text{data}}\boldsymbol{L}(\boldsymbol{G}(t))+x\sim p_{\text{data}}\boldsymbol{L}(\boldsymbol{G}(t))}{2}\right)-2\ln 2\\&=2\cdot\mathrm{JS}(x\sim p_{\text{data}}\boldsymbol{G}(t)\|x\sim p_{\text{data}}\boldsymbol{L}(\boldsymbol{G}(t)))-2\ln 2\end{aligned} \tag{17}$$

上述判别器D通过衡量特征信息矩阵和预测结果矩阵之间的差异，以此计算多指标预测的综合误差，在训练过程中不断找出小麦指标预测中误差较大的数值，并传递给LSTM网络进一步优化，LSTM-GAN模型也会逐渐降低真实数据分布$x\sim P_{g(t)}$和预测结果分布$x\sim P_{l(g(t))}$之间的JS散度，使得预测结果的综合误差逐渐减小，因此，LSTM-GAN模型可通过对抗训练的方法提取出小麦多指标数据序列的变化特征，逐渐匹配到小麦真实指标数据的特征分布，使多指标的预测结果更加准确。

3.3 LSTM-GAN模型优化

LSTM-GAN模型的训练过程可采用反向传播算法和实时递归算法[10]，通过调整相应的权值信息使结果误差逐渐降低，其中反向传播算法效率更高，在计算时间上具有优势。本文改进的LSTM-GAN模型中输入门i_t和输出门o_t作用不变，为了使LSTM记忆单元存储更为有效的时序信息，其中遗忘门f_t根据判别器D的计算结果选择性的遗忘掉无效的LSTM记忆单元信息，可采用如式(18)和式(19)的计算过程

$$f_t = \sigma(+\text{loss}\ g_t) \cdot (\boldsymbol{W}_{xf}x_t + \boldsymbol{W}_{hf}h_{t-1} + \boldsymbol{W}_{cf}c_{t-1} + \boldsymbol{b}_f) \tag{18}$$

$$\text{loss} = \frac{1}{2}\sum_{k=1}^{m}(y_k - y_k^*)^2 \tag{19}$$

其中，y_k^*为第k个神经元的期望输出，m为输出层神经元的个数，最后根据相应的误差项，计算每个权重的梯度，反复应用链式规则使整体损失函数达到最小。用q代表训练过程的迭代次数，则第k个神经元权值的更新公式为

$$W^{q+1} = W^q + \Delta W^q = Wq - \eta\frac{\partial \text{loss}}{\partial W}\bigg|_{W=W^q} \tag{20}$$

由于小麦储藏期间多指标随储藏时间变化的数值变化幅度存在一定差异，所以小麦品质受不同时期储藏条件影响导致的劣变程度不一，因此，采用LSTM对小麦多指标数据进行时序建模时，可选择不同的训练窗口长度进行训练并计算得出最优的超参数，以区别不同储藏时期对于品质的影响程度，再通过以上反向误差传播算法训练LSTM-GAN模型得出多指标预测结果。

4 实验结果分析

为了检验LSTM-GAN模型预测小麦品质多指标数据的准确性，本文选用文献[11]中3种筋力(强筋、中筋、弱筋)及4种环境温度(25 °C, 30 °C, 35 °C和40 °C)的仓储小麦在0～210 d储藏时间内测试间隔为30 d的6个生理生化指标的测试数据。根据交叉验证法把96(3种筋力×4种环境温度×8个采样点)组时序数值划分成72组训练集和24组测试集，其中每组时序数值均包含相应的6个小麦生理生化指标数值(脂肪酸值、降落数值、沉降值、发芽率、过氧化物酶、电导率)，把它们分别输入到LSTM-GAN模型进行训练和预测分析，其中取30 °C, 35 °C和40 °C这3种温度下指标数据作为训练集，预测分析25 °C下小麦指标测试集的变化规律。

所选用数据特征信息如表1所示。由表1可知多指标数值范围差异较大，为避免因多指标数值差异及计量单位对模型训练误差所产生的影响，采用Z-score公式[12]对多指标时序数据进行预处理，为后续的小麦品质预测分析提供标准化的指标数据集。

4.1 LSTM-GAN模型结构组合优化

经大量实验[13]表明，LSTM网络的结构参数通常可对训练结果产生较大影响，因此本文LSTM-GAN模型将主要分析训练时序窗口长度、隐含层层数和其中神经元个数对于模型训练的效率及准确度所产生的影响。采用相同的LSTM-GAN模型结构参数(隐含层神经元个数为10，隐含层层数为2)，来讨论不同的训练时序窗口长度这一变量对多指标数据进行训练及预测的影响，在LSTM-GAN模型中梯度下降采用Adam方法优化学习率，计算训练时序窗口长度为2, 4, 6和8时各个指标的预测误差数据如表2所示。

表2中，当训练窗口时序长度为4时，脂肪酸值、电导率取得了较小的误差；而训练窗口时序长度为6时，降落数值、沉降值、发芽率、过氧化物酶指标的训练误差较小，因此，不同指标对小麦品质变化趋势的长短期记忆信息的依赖程度不同，综合多指标可知当训练窗口长度为6时整体预测误差最小。

除训练窗口长度外，LSTM-GAN模型中隐含层的网络结构仍需进一步优化，在模型时序长度为6时，采用2, 3和5层隐含层和每层6, 8, 10和12个神经元对多指标数据进行训练，计算得出的模型训练误差数据如表3所示。

由表3可知，在6个指标中发芽率的预测误差最

表 1　小麦多指标数据集统计信息

	最小值	最大值	均值	标准差
脂肪酸值(mgKOH/100 g)	16.00	30.50	23.18	4.24
降落数值(s)	365.00	630.00	482.81	69.36
沉降值(ml)	19.50	62.00	40.11	13.94
发芽率(%)	0	97.00	71.29	28.96
过氧化物酶(U/g)	1400.00	4100.00	3171.35	667.93
电导率(μs/(cm·g))	25.50	60.50	39.11	8.75

表 2　模型不同训练窗口长度误差对比

窗口长度	2	4	6	8
脂肪酸值	0.260	0.258	0.308	0.328
降落数值	0.325	0.263	0.228	0.277
沉降值	0.356	0.447	0.336	0.407
发芽率	0.652	0.530	0.483	0.511
过氧化物酶	0.424	0.455	0.402	0.415
电导率	0.412	0.324	0.329	0.374

大，这是由于在特定条件下该指标预测值偏离了模型拟合的时序数据，进而造成了更大的误差。表3还显示实验中隐含层层数越多，误差也随之逐渐增大，这是由于隐含层层数过多产生了过拟合现象，2层隐含层的网络结构已经能够得出较低的预测误差结果；同时隐含层神经元个数并不是越多越好，需要根据数据分布选择适当的神经元个数，当隐含层神经元个数为10时有助于快速训练得出更准确的预测结果。

4.2 LSTM-GAN模型预测结果分析

通过对模型结构参数的优化，可得到LSTM-GAN预测多指标时序数据的综合误差，仍需要进一步对比分析小麦指标在筋力方面的表现，为此，以强筋麦小麦为例，比较多指标真实值与预测值如图3所示。由图3可知，在某些特定条件下，如发芽率指标的第6～8批次、过氧化物酶指标的第3批次，对应的预测值误差较大，说明小麦品质在此期间发生了过快劣变导致实际值低于预测值，因此在多指标预测中应充分考虑储藏时期及环境变化对多指标数据造成的影响。

另外为比较不同筋力小麦的指标预测情况，用LSTM-GAN模型分别训练强筋、中筋、弱筋这3种筋力小麦多指标数据，得出如表4所示的误差计算结果。其中，脂肪酸值、电导率这两个指标预测强筋麦的数据误差相对较小；对于中筋麦，发芽率、过氧化物酶这2个指标的预测更加准确；降落数

表 3 LSTM-GAN模型不同结构参数训练误差

隐含层层数	2				3				5			
神经元个数	6	8	10	12	6	8	10	12	6	8	10	12
脂肪酸值	0.285	0.245	0.275	0.281	0.265	0.290	0.260	0.285	0.255	0.355	0.345	0.335
降落数值	0.295	0.265	0.305	0.335	0.315	0.235	0.300	0.342	0.335	0.315	0.335	0.355
沉降值	0.400	0.405	0.410	0.427	0.405	0.425	0.435	0.533	0.445	0.540	0.315	0.493
发芽率	0.505	0.560	0.488	0.494	0.610	0.570	0.532	0.582	0.635	0.623	0.657	0.625
过氧化物酶	0.365	0.345	0.340	0.342	0.370	0.280	0.300	0.369	0.325	0.380	0.415	0.409
电导率	0.330	0.370	0.340	0.404	0.440	0.375	0.425	0.417	0.555	0.370	0.435	0.454
综合误差	2.180	2.190	2.158	2.284	2.405	2.175	2.252	2.528	2.550	2.583	2.502	2.671

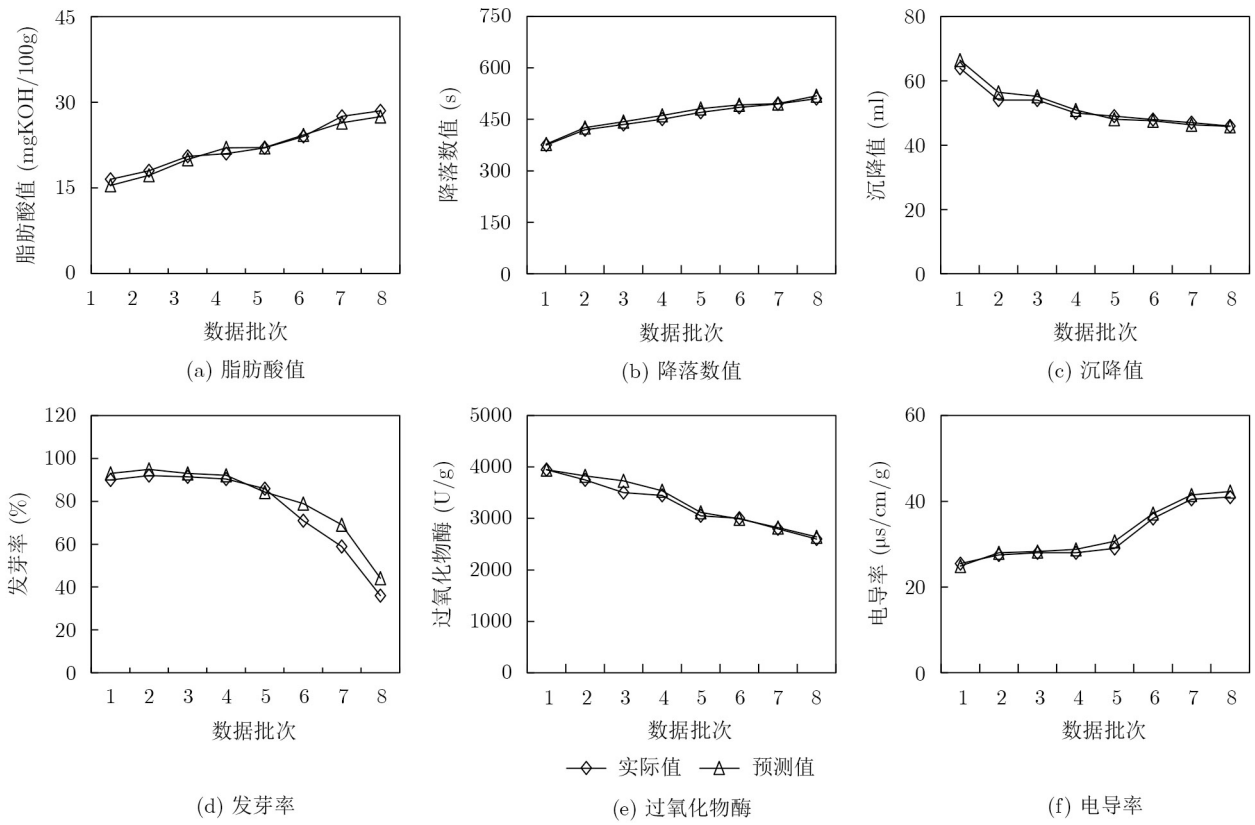

图 3 强筋麦多指标预测结果

值、沉降值指标在弱筋麦中的预测误差小于强筋麦和中筋麦的预测结果。由表4可知，3种筋力小麦的总体误差分别为2.042, 1.974和1.943，它们没有明显的数值差异。

4.3 LSTM-GAN与其他模型预测误差对比

为比较本文模型相对于其他时序预测模型的准确度差异，使用相同的小麦多指标时序数据，选择LSTM、多元线性回归模型[14]、支持向量机回归模型(Support Vector Regression, SVR)[15]、人工神经网络(Artificial Neural Network, ANN)[16]、灰色预测模型(Grey Model, GM)[17]这5种常用预测模型的计算结果作为对比，通过对模型的参数进行优化可得到不同模型的预测结果如表5所示。其中SVR模型采用径向基核函数；LSTM-GAN模型根据参数优化结果取时序长度为6、隐含层为2、每层神经元个数为10；LSTM, SVR, ANN和LSTM-GAN这4种模型采用相同的训练次数。

由表5可知，LSTM-GAN模型预测多指标取得了最小的综合误差2.158，它比LSTM模型的综合误差2.391降低了0.233，准确度提升了9.745%。这是由于本文模型以多指标的对抗学习预测小麦品质的整体变化趋势，并结合LSTM来降低多指标序列综合误差，而经典的LSTM算法中没有涉及多目标的组合优化，导致了预测结果仍存在较大误差，因此，本文LSTM-GAN模型更适用于小麦多指标预测分析。

对比不同模型的多生理生化指标预测结果发现发芽率指标的预测误差明显高于其他指标，如果去除发芽率指标后再用LSTM, LSTM-GAN, 线性回归, SVR, ANN和GM预测其余5个指标，则预测误差分别下降了0.553, 0.488, 0.611, 0.467, 0.466和0.559，对应的误差下降百分比分别为13.760%, 22.602%, 20.531%, 8.505%, 18.756%和19.837%，因此，为提升小麦品质多指标预测综合结果的准确性，将发芽率和其他指标分开进行预测分析是更佳的选择。

表 4　不同筋力小麦多指标预测误差对比

	强筋	中筋	弱筋
脂肪酸值	0.275	0.295	0.315
降落数值	0.305	0.290	0.255
沉降值	0.360	0.320	0.245
发芽率	0.422	0.419	0.428
过氧化物酶	0.390	0.350	0.365
电导率	0.290	0.300	0.335

表 5　不同模型预测误差对比

	LSTM-GAN	LSTM	线性回归	SVR	ANN	GM
脂肪酸值	0.275	0.285	0.290	0.303	0.326	0.386
降落数值	0.305	0.329	0.577	0.405	0.402	0.511
沉降值	0.410	0.482	0.563	0.366	0.459	0.498
发芽率	0.488	0.553	0.611	0.467	0.466	0.559
过氧化物酶	0.340	0.378	0.604	0.469	0.460	0.452
电导率	0.340	0.364	0.331	0.372	0.373	0.413
综合误差	2.158	2.391	2.976	2.381	2.484	2.817

5　结束语

针对小麦品质指标预测中因数值范围、计量单位差异等因素而产生较大综合误差的问题，本文提出了一种改进拓扑结构的LSTM-GAN模型用于多指标预测分析。实验分析表明：适当的训练时序长度及隐含层结构参数均有助于降低预测结果的误差，小麦多指标数据对于长短期记忆信息的依赖程度存在一定差异，因此需要根据多指标序列综合误差组合优化模型的训练窗口长度及网络结构；另外某些特定条件下发芽率、过氧化物酶指标的误差较高，对比预测值和实际值发现小麦品质在此期间发生了过快劣变，因此在多指标预测中应充分考虑储藏时期及环境变化对多指标数据造成的影响；经对比分析不同模型的预测误差，改进的LSTM-GAN模型比LSTM预测多指标的整体误差下降了9.745%，其综合误差低于线性回归, SVR, ANN和GM预测模型，可有效提高小麦多指标预测结果的准确性。

参 考 文 献

[1] KALSA K K, SUBRAMANYAM B, DEMISSIE G, *et al*. Evaluation of postharvest preservation strategies for stored wheat seed in Ethiopia[J]. *Journal of Stored Products Research*, 2019, 81: 53–61. doi: 10.1016/j.jspr.2019.01.001.

[2] ZHANG Shuaibing, LÜ Yangyong, WANG Yuli, *et al*. Physiochemical changes in wheat of different hardnesses during storage[J]. *Journal of Stored Products Research*, 2017, 72: 161–165. doi: 10.1016/j.jspr.2017.05.002.

[3] 陈红松, 陈京九. 基于循环神经网络的无线网络入侵检测分类模型构建与优化研究[J]. 电子与信息学报, 2019, 41(6): 1427–1433. doi: 10.11999/JEIT180691.
CHEN Hongsong and CHEN Jingjiu. Recurrent neural networks based wireless network intrusion detection and classification model construction and optimization[J]. *Journal of Electronics & Information Technology*, 2019, 41(6): 1427–1433. doi: 10.11999/JEIT180691.

[4] XU Peng, DU Rui, ZHANG Zhongbao, *et al*. Predicting pipeline leakage in petrochemical system through GAN and LSTM[J]. *Knowledge-Based Systems*, 2019, 175: 50–61. doi:

10.1016/j.knosys.2019.03.013.

[5] MAHASSENI B, LAM M, and TODOROVIC S. Unsupervised video summarization with adversarial lstm networks[C]. 2017 IEEE conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 2982–2991. doi: 10.1109/CVPR.2017.318.

[6] YANG Yang, ZHOU Jie, AI Jiangbo, *et al*. Video captioning by adversarial LSTM[J]. *IEEE Transactions on Image Processing*, 2018, 27(11): 5600–5611. doi: 10.1109/TIP.2018.2855422.

[7] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, *et al*. Generative adversarial nets[C]. The 27th International Conference on Neural Information Processing Systems, Montreal, Canada, 2014: 2672–2680.

[8] 曹志义, 牛少彰, 张继威. 基于半监督学习生成对抗网络的人脸还原算法研究[J]. 电子与信息学报, 2018, 40(2): 323–330. doi: 10.11999/JEIT170357.

CAO Zhiyi, NIU Shaozhang, and ZHANG Jiwei. Research on face reduction algorithm based on generative adversarial nets with semi-supervised learning[J]. *Journal of Electronics & Information Technology*, 2018, 40(2): 323–330. doi: 10.11999/JEIT170357.

[9] 蒋华伟, 张磊, 周同星. 基于信息熵的小麦储藏品质多指标权重模型研究[J]. 中国粮油学报, 2020, 35(6): 105–113. doi: 10.3969/j.issn.1003-0174.2020.06.016.

JIANG Huawei, ZHANG Lei, and ZHOU Tongxing. Research on multi-index weight model of wheat storage quality based on information entropy[J]. *Journal of the Chinese Cereals and Oils Association*, 2020, 35(6): 105–113. doi: 10.3969/j.issn.1003-0174.2020.06.016.

[10] 刘威, 刘尚, 白润才, 等. 互学习神经网络训练方法研究[J]. 计算机学报, 2017, 40(6): 1291–1308. doi: 10.11897/SP.J.1016.2017.01291.

LIU Wei, LIU Shang, BAI Runcai, *et al*. Research of mutual learning neural network training method[J]. *Chinese Journal of Computers*, 2017, 40(6): 1291–1308. doi: 10.11897/SP.J.1016.2017.01291.

[11] 高艳娜. 小麦产后品质变化规律研究[D]. [硕士论文], 河南工业大学, 2010.

GAO Yanna. Study on the changes of postpartum quality in wheat[D]. [Master dissertation], Henan University of Technology, 2010.

[12] FRIEDMAN L and KOMOGORTSEV O V. Assessment of the effectiveness of seven biometric feature normalization techniques[J]. *IEEE Transactions on Information Forensics and Security*, 2019, 14(10): 2528–2536. doi: 10.1109/TIFS.2019.2904844.

[13] GREFF K, SRIVASTAVA R K, KOUTNÍK J, *et al*. LSTM: A search space odyssey[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(10): 2222–2232. doi: 10.1109/TNNLS.2016.2582924.

[14] FANG Tingting and LAHDELMA R. Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system[J]. *Applied Energy*, 2016, 179: 544–552. doi: 10.1016/j.apenergy.2016.06.133.

[15] XU Jie, XU Chen, ZOU Bin, *et al*. New incremental learning algorithm with support vector machines[J]. *IEEE Transactions on Systems, Man, and Cybernetics; Systems*, 2019, 49(11): 2230–2241. doi: 10.1109/tsmc.2018.2791511.

[16] VILLARRUBIA G, DE PAZ J F, CHAMOSO P, *et al*. Artificial neural networks used in optimization problems[J]. *Neurocomputing*, 2018, 272: 10–16. doi: 10.1016/j.neucom.2017.04.075.

[17] DING Song, HIPEL K W, and DANG Yaoguo. Forecasting China's electricity consumption using a new grey prediction model[J]. *Energy*, 2018, 149: 314–328. doi: 10.1016/j.energy.2018.01.169.

蒋华伟：男，1970年生，博士，教授，博士生导师，研究方向为粮食信息处理.

张 磊：男，1996年生，硕士生，研究方向为粮食多指标智能预测.

责任编辑：马秀强