

# 基于 CNN-BiLSTM-AM 的储层岩石脆性指数预测

杜睿山<sup>1,2</sup>, 李宏杰<sup>1</sup>, 孟令东<sup>2</sup>

(1. 东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318;

2. 油气藏及地下储库完整性评价黑龙江省重点实验室, 黑龙江 大庆 163318)

**摘要:** 脆性指数是储层岩石的重要地质力学性质之一, 但对于脆性指数至今为止都没有一个明确的定义, 许多学者提出了不同的方法来测量该参数, 一些方法如矿物分析等成本较高, 然而机器学习和深度学习可以有效融合多元数据, 充分利用数据去挖掘自变量与因变量之间的关系, 且成本较低。因此, 该文使用深度学习方法来构建测井曲线数据与储层岩石脆性之间的关系模型。因测井曲线是特殊的时序曲线, 该文采用可以综合考虑过去和未来信息的 BiLSTM (双向长短期记忆) 模型, 同时为了进一步的优化, 在模型中添加 1DCNN (一维卷积) 用来提取特征, 且引入了 AM (注意力机制)。同时利用 Pearson 系数和 XGBoost (极限梯度提升树) 进行分析, 研究了各测井曲线对脆性的敏感性等级以及重要性程度, 最终选取的输入参数有 AC (声波时差)、DEN (密度)、CAL (井径)、GR (伽马射线) 和 SP (自然电位)。与其它机器学习方法相比, 该方法的 MSE 和 RMSE 最小, 分别为 0.003 5 和 0.059 16, 表明 CNN-BiLSTM-AM 是一种预测精度更高、效果更好的方法。

**关键词:** 测井曲线; 脆性指数; 深度学习; Pearson 系数; BiLSTM; 一维卷积; 注意力机制

中图分类号: TP391.1

文献标识码: A

文章编号: 1673-629X(2023)10-0028-07

doi: 10.3969/j.issn.1673-629X.2023.10.005

## Prediction of Reservoir Rock Brittleness Index Based on CNN-BiLSTM-AM

DU Rui-shan<sup>1,2</sup>, LI Hong-jie<sup>1</sup>, MENG Ling-dong<sup>2</sup>

(1. School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China;

2. Key Laboratory of Oil and Gas Reservoir and Underground Gas Storage Integrity Evaluation, Daqing 163318, China)

**Abstract:** Brittleness index is one of the important geomechanical properties of reservoir rocks, but there is no clear definition for brittleness index so far. Many scholars had proposed different methods to measure this parameter. Some methods, such as mineral analysis, had higher costs. However, machine learning and depth learning can effectively integrate multivariate data, and make full use of the data to mine the relationship between independent variables and dependent variables with lower costs. Therefore, we used depth learning method to build the relationship model between logging curve data and reservoir rock brittleness. Because the logging curve was a special time series curve, we adopted the BiLSTM (Bi-directional Long Short-Term Memory) model that can comprehensively consider past and future information. At the same time, for further optimization, 1DCNN (One-Dimensional Convolution) was added to the model to extract features, and AM (Attention Mechanism) was introduced. At the same time, Pearson coefficient and XGBoost (eXtreme Gradient Boosting) were used for analysis, and the sensitivity level and importance of each logging curve to brittleness are studied. The final selected input parameters were AC (Acoustic moveout), DEN (Density), CAL (Caliper), GR (Gamma Ray) and SP (Spontaneous Potential). Compared with other machine learning methods, the proposed method has the smallest MSE and RMSE, 0.003 5 and 0.059 16 respectively. It is showed that CNN-BiLSTM-AM is a method with higher prediction accuracy and better effect.

**Key words:** logging curve; brittleness index; deep learning; Pearson coefficient; BiLSTM; one dimensional convolution; attention mechanism

收稿日期: 2022-11-17

修回日期: 2023-03-20

基金项目: 国家自然科学基金青年科学基金(41702156); 东北石油大学引导性创新基金(2020YDL-04)

作者简介: 杜睿山(1977-), 男, 通讯作者, 硕士, 副教授, CCF 会员(51501M), 研究方向为人工智能、机器学习等; 李宏杰(1997-), 男, 硕士研究生, 研究方向为机器学习等。

## 0 引言

在处理岩石工程时,脆性被认为是一个重要的岩石参数。例如,在天然气和石油工程领域,评估井筒稳定性以及评估水力压裂作业的性能需要充分了解岩石脆性。岩石脆性是指岩石在发生很小变形即破坏的性质,是井筒稳定性评价和水力压裂评估的参数之一<sup>[1]</sup>。在非常规油气勘探开发过程,特别是页岩和致密砂岩油气,储层脆性评价和预测更是不可或缺的研究内容<sup>[2]</sup>。

脆性的定义已经提出了很多,但由于工程中需要的目的不同,脆性的测量和识别仍然没有统一的规则。目前测井数据评价岩石脆性的方法可以分为以下几类:基于脆性矿物含量的脆性评价、基于岩石力学参数的脆性评价和基于应力-应变曲线的脆性指数评价。现在应用较为广泛的是基于岩石力学参数的脆性指数评价,其计算需要杨氏模量和泊松比等弹性参数,而参数的计算需要横波测井和岩性-密度测井,由于测井服务成本高,传统上不采用,因此有时需要从局部经验拟合方程粗略计算弹性参数,这可能会给最终的脆性结果带来较大的误差。另外,由于岩石的脆性特征是在一定的地层条件下的岩石不同矿物组成、结构和物理力学性质的综合结果<sup>[3]</sup>。因此,想要全面体现岩石的脆性特征,需要综合考虑多条测井曲线。

近年来,岩石脆性已被许多地质力学领域的研究者广泛研究,大量研究使用机器学习(ML)和人工智能(AI)技术来解决科学和工程领域的相关问题。一些学者也比较倾向于建立基于数据的脆性关系,具体的测井曲线或储层属性与脆性指数的关系。部分学者的研究中,已证实了脆性指数与一些常规测井曲线具有良好的关系。在一些页岩气储层中,矿物学脆性与自然伽马/光电吸收截面指数(自然伽马与光电吸收截面指数比)、孔隙度、纵波时差等之间存在良好的关系<sup>[4-6]</sup>。时贤等人<sup>[7]</sup>利用人工智能技术预测了岩石脆性,并在论文中说明矿物学信息和测井信息的结合有助于更好地扩展有限的矿物学数据,同时反映井下情况。袁思乔等人<sup>[8]</sup>基于多测井曲线利用层次分析法对脆性指数进行了预测。叶亚培等人<sup>[9]</sup>基于常规测井数据,利用 PCA-BPNN 的方法对脆性指数进行了预测,实验结果表明,该方法具有较高的准确性。

## 1 相关工作

根据前人的研究表明,利用常规测井曲线去预测脆性指数是可行的,因此,该文利用机器学习和深度学习的方法进行脆性预测,基于岩石力学参数杨氏模量和泊松比构建的脆性数据集,建立常规测井参数与脆性指数的关系模型,准确地将脆性指数与测井参数关

联起来。这种关联的优势在于:可以使用廉价且相对简单的预测技术来估计整口井的脆性指数,而且大多数井的测井输入数据都可以在整口井中获得,而不需要收集额外数据。

之前的一些研究承认机器学习技术在解决工程问题方面的适用性,但仍有一些技术未被使用或几乎未被用于预测岩石脆性指数<sup>[10]</sup>。前人的研究几乎都是利用传统机器学习方法或者统计学习方法,虽然可以建立常规测井曲线与脆性指数的模型,但是未能考虑测井曲线的序列特征,因此可能会带来一定的误差。该文选取双向长短期记忆神经网络,因为其可以综合考虑过去和未来信息,使得序列信息在模型中一直传递下去,从而更能保证结果的有效性。

该文基于常规测井曲线数据,选取多条测井曲线数据为输入,以脆性指数为输出,首次将双向长短期记忆神经网络引入到脆性指数预测中,采用 CNN-BiLSTM-AM 神经网络进行实验,构建出常规测井曲线和脆性指数的关系模型,对储层岩石的脆性指数进行预测与分析,并和多种机器学习方法进行对比,以验证该模型的优越性。借助 Pearson 相关系数分析,研究了各测井曲线对脆性的敏感性等级。利用这些变量,比较了若干种机器学习模型和文中所提模型的性能。从神经网络模型获得的脆性数值结果揭示了 CNN-BiLSTM-AM 在预测过程中的高精度和高效率,这有助于使用现有的实验数据进行更好的脆性预测。

## 2 方法原理

### 2.1 CNN 概述

CNN(卷积神经网络)的局部感知和权值共享可以大大减少参数的数量,从而提高模型的学习效率。CNN 主要由三部分组成:卷积层、池化层和全连接层<sup>[11]</sup>。每个卷积层包含多个卷积核,其计算公式见公式(1)。对卷积层进行卷积运算后,提取数据的特征。然而,提取的特征维数非常高。所以为了解决这个问题,降低网络训练的成本,在卷积层之后增加了一个 pooling 层来降低特征维数。

$$l_i = \tanh(x_i * k_i) + b_i \quad (1)$$

其中, $l_i$  为卷积后的输出值, $\tanh$  为激活函数, $x_i$  为输入向量, $k_i$  为卷积核的权值, $b_i$  为卷积核的偏置值。

### 2.2 LSTM 与 BiLSTM 概述

#### 2.2.1 LSTM

常规的 RNN 会发生梯度消失或者梯度爆炸问题<sup>[12]</sup>,导致不收敛,从而致使网络预测效果不佳。为此,后来的学者对其进行了改良,LSTM 通过添加 3 个门来控制不同时期对后续信息的影响同时使用隐状态和细胞状态传递信息,使得 LSTM 能够综合局部信息

和序列信息<sup>[13]</sup>。

图1是常见的长短期模型记忆单元,主要由遗忘门、输入门和输出门这三个门组成。

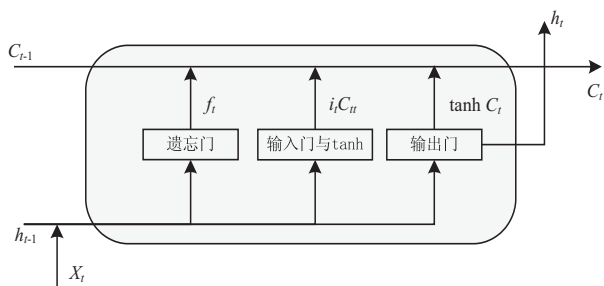


图1 LSTM网络重复单元结构

每个门的具体公式如下所示：

$$f_t = \sigma[W_f(h_{t-1}, x_t) + b_f] \quad (2)$$

$$i_t = \sigma[W_i(h_{t-1}, x_t) + b_i] \quad (3)$$

$$C_t = f_t C_{t-1} + i_t C_u \quad (4)$$

$$O_t = \sigma[W_o(h_{t-1}, x_t) + b_o] \tanh C_t \quad (5)$$

式中,  $f_t$  是  $t$  时刻遗忘门的输出,  $h_{t-1}$  为前一时刻的隐状态,  $x_t$  为当前时刻的输入,  $O_t$  是  $t$  时刻输入门的输出;  $C_u$  是  $t$  时刻  $\tanh$  层的输出;  $C_t$  为  $t$  时刻的细胞状态,  $i_t$  是  $t$  时刻输入门的输出,  $W_f$ 、 $b_f$  是遗忘门的权重和偏置;  $W_i$ 、 $b_i$  是输入门的权重和偏置;  $W_o$ 、 $b_o$  是输出门的权重和偏置。

### 2.2.2 BiLSTM

BiLSTM(双向长短期记忆)神经网络由两层独立的LSTM层构成(如图2所示),一层前向LSTM和一层后向LSTM,是对LSTM的改进,BiLSTM可以综合考虑历史信息与未来信息,所以较普通LSTM有更好的效果。

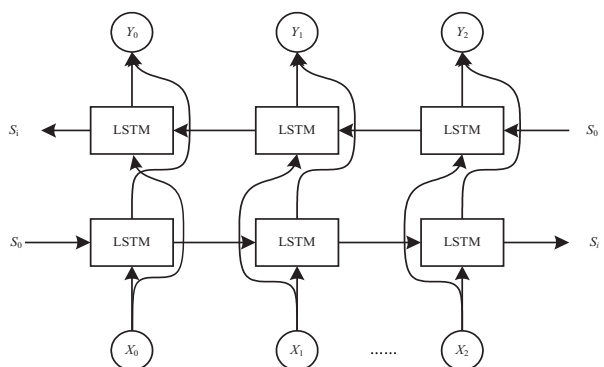


图2 BiLSTM示例图

### 2.3 AM概述

注意力机制(Attention Mechanism)的主要思想来源于人类视觉注意的过程<sup>[14]</sup>。AM选择性地关注一些比较重要的信息,忽略不重要的信息,可以分配信息的重要性。

如图3所示,AM的计算过程一般分为三个阶段:(1)相似度计算阶段;(2)归一化处理阶段;(3)注意力

加权求和阶段,每阶段具体公式如下:

$$S_t = \tanh(W_h h_t + b_h) \quad (6)$$

$$a_t = \frac{\exp(s_t^T v)}{\sum_i \exp(s_i^T v)} \quad (7)$$

$$s = \sum_i a_i h_t \quad (8)$$

其中,  $W_h$  为AM的权值,  $b_h$  为AM的偏置,  $h_t$  为输入向量,  $W_h$ 、 $b_h$  为各层的共享权值,  $v$  为注意值。

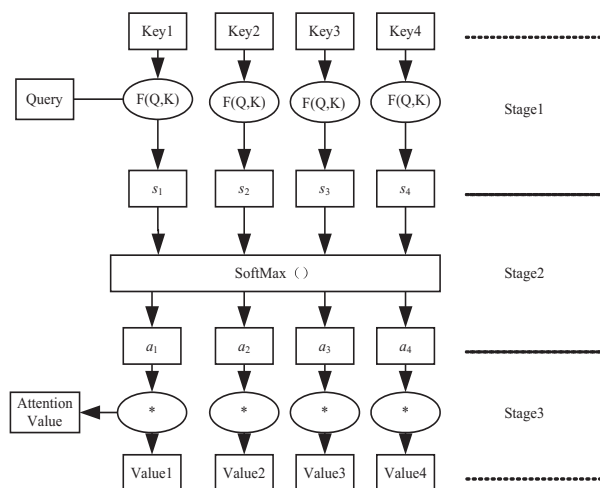


图3 AM流程

### 2.4 CNN-BiLSTM-AM模型

CNN具有关注视线中最明显的特征的特点,因此在特征工程中被广泛应用。BiLSTM具有按时间序列展开的特点,在时间序列分析中应用广泛。AM具有将时间序列数据的过去特征状态添加到输出结果的重要特性。

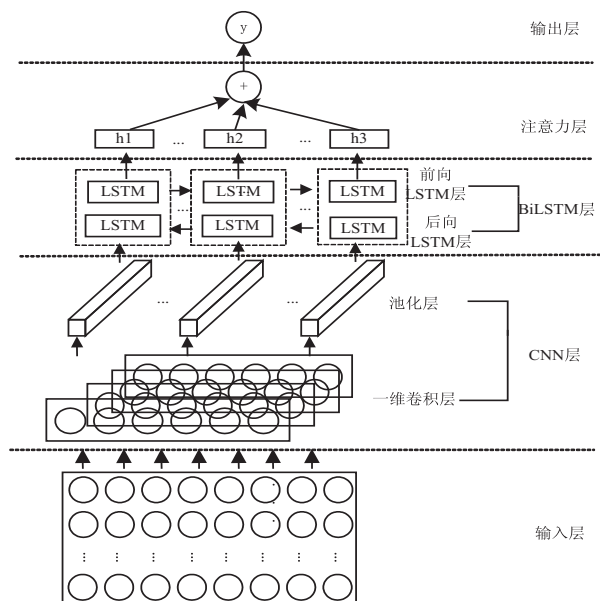


图4 CNN-LSTM-AM模型

根据CNN、BiLSTM和AM的特点,充分发挥各自的优势,该文提出了一种基于CNN-BiLSTM-AM的岩石脆性预测模型。模型结构如图4所示,主要结构

有 CNN、BiLSTM、AM,包括输入层、CNN 层(一维卷积层、池化层)、BiLSTM 层(正向 LSTM 层、反向 LSTM 层)、AM 层、输出层。

(1)输入层:对脆性数据集进行切片处理,并将处理好的数据作为 CNN 层的输入。

(2)CNN 层:其有共享权重和局部感知等特点,这意味着 CNN 具有独特的优势。该文使用卷积和池化操作从测井曲线序列数据中提取特征。

(3)BiLSTM 层:该层由正向 LSTM 和逆向 LSTM 组成。与 LSTM 相比,BiLSTM 可以双向提取测井曲线序列信息,更好地收集来自 CNN 的数据的正向和反向依赖关系,对特征信息进行选择性记忆和遗忘,学习测井信息随深度的变化趋势和前后相关性,有效防止梯度消失或爆炸。

(4)AM 层:注意力层可以对特征进行进一步的信息筛选,对不同的特征赋予不同的权重,有选择地从模型训练过程中学习特征和关键信息,过滤无用信息,从而提高模型预测的准确性。

(5)输出层:输出 AM 层加权求和的预测结果,然后将测试集输入训练好的模型进行预测,生成脆性指数预测结果。

该文研究的一种基于 CNN-BiLSTM-AM 的脆性指数预测模型具体流程如图 5 所示。

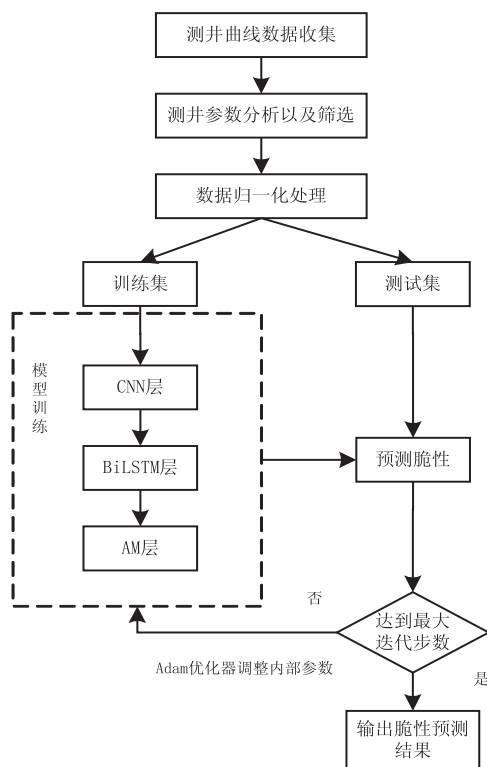


图 5 CNN-LSTM-AM 流程

(1)采集区域内井的数据,构建出基于岩石弹性参数的脆性指数数据集。

(2)对测井数据进行处理,并且进行 Pearson 线性

相关性分析以及 XGBoost 特征重要性分析,筛选特征,提升模型预测的精度以及速度。

(3)划分步骤二中归一化后的数据集,按照 7:3 的比例划分为训练集和测试集。

(4)利用前馈传播计算输出本次迭代的预估值,然后再反向计算每个神经元的误差项值,经过 Adam 优化参数后继续训练模型。

(5)将训练步数作为终止条件,若达到终止条件,则返回最优结果,否则继续步骤四。

### 3 实验及分析

实验环境:实验在 Windows 平台进行,计算机具体配置如下:处理器为酷睿 i7-9700K,显卡为英伟达 GTX2070 和 16G 内存条,所提模型基于 python 深度学习框架。

#### 3.1 数据获取

实验数据集来源于胜利油田某井的常规测井数据,脆性指数是基于岩石力学参数杨氏模量和泊松比计算而来的,具体计算公式如式(9)所示:

$$BI_0 = \frac{\Delta E_s + \Delta PR_s}{2} \quad (9)$$

其中:

$$\Delta E_s = \frac{E_s - 1}{8 - 1} \quad (10)$$

$$\Delta PR_s = \frac{0.4 - PR_s}{0.4 - 0.15} \quad (11)$$

式中,  $BI_0$  为岩石脆性指数;  $\Delta E_s$  为杨氏模量;  $\Delta PR_s$  为泊松比;  $E_s$  为静态杨氏模量  $10^6 \text{ psi}$  ( $1 \text{ psi} = 0.0068948 \text{ MPa}$ );  $PR_s$  为静态泊松比,无量纲。

#### 3.2 测井参数分析

该文收集了完整的测井数据集,包括声波时差(AC)、井径(CAL)、密度(DEN)、中子(CNL)、伽马射线(GR)、自然电位(SP)和深度侧向测井(LLD)等曲线。其中 LLD 测井曲线表示地层电阻率,主要受地层流体和饱和度控制,与岩石力学性质无关,因此首先剔除 LLD 测井曲线。其次,通过分析实测 BI(脆性指数)值与相应的测井数据之间的关系,选择合适的测井参数进行脆性评价是模型构建的关键步骤,因为输入参数与脆性指数值密切相关的测井数据可以准确地预测脆性指数值。以 Pearson 相关系数为关键指标,考察脆性对测井曲线的线性依赖程度,具体的 Pearson 相关系数计算公式如下:

$$\rho(X, Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sqrt{\sum_{i=1}^n (X_i - \mu_X)^2} \sqrt{\sum_{i=1}^n (Y_i - \mu_Y)^2}} \quad (12)$$



其中,  $E[(X - \mu_X)(Y - \mu_Y)]$  称为随机变量  $X$  与  $Y$  的协方差,  $\sqrt{\sum_{i=1}^n (X_i - \mu_X)^2}$  和  $\sqrt{\sum_{i=1}^n (Y_i - \mu_Y)^2}$  是两个随机变量之间的标准差;把两个变量之间的协方差和标准差的商称为随机变量  $X$  与  $Y$  的相关系数,记为  $\rho(X, Y)$ ;Pearson 相关系数是衡量两个变量之间相关程度的一种方法, Pearson 相关系数是在 -1 到 1 之间的数值,其取值越大,则代表两个变量之间的相关性越强。

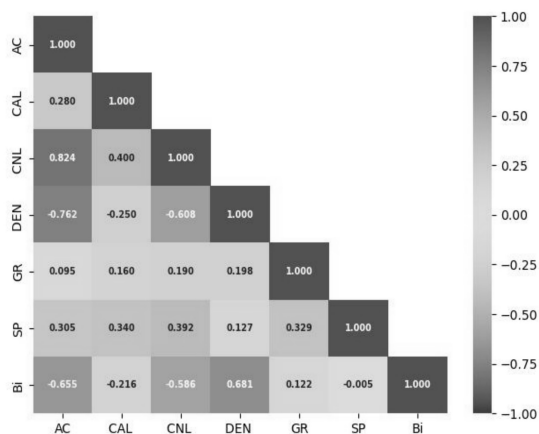


图6 皮尔逊相关系数

从图中6可以看出,除了SP与BI无明显的线性相关性,其余曲线都与BI有着或多或少的相关性。AC、CNL、CAL与BI呈负相关性,且AC、CNL与BI的负相关性较大;DEN、GR与BI成正相关,其中部分结果可以从岩石物理角度加以说明。

AC、CNL、DEN曲线与脆性指数有着重要的关系,因为岩石孔隙度的增加会导致岩石脆性的降低。总有机碳(TOC)含量高的页岩会产生较高的伽马射线值,前人研究表明,有机质主要富集在硅质页岩和碳质页岩中,由于存在适量的石英和碳酸盐矿物,这类岩石具有较高的BI值,因此,BI高的岩石通常具有较高的GR<sup>[9]</sup>。本数据集中,SP曲线虽然与BI无明显的线性相关性,但SP与岩石孔隙度相关,也是一个十分重要的参数。从以上可以看出,本研究所选取的测井曲线几乎都是与BI有着重要的物理关系。

### 3.3 特征重要性分析以及特征筛选

虽然CAL、GR、SP与BI的线性相关性并不明显,但是不排除它们之间有着非线性关系,因此有必要进一步分析。采用XGBoost方法进行特征重要性分析,XGBoost可以找到特征和目标之间的非线性关系,其在分裂时会使用最高信息增益的特征作为分裂条件。当某一个特征多次被用来构建决策树,其重要性就会越高,经XGBoost排序后的重要性如表1所示。

考虑到特征之间可能存在线性相关性较高的特征,即冗余特征,因此对两两测井参数之间进行

Pearson相关性分析,图6可以看到测井曲线之间的线性关系。AC与CNL曲线相关性程度高达0.824,根据Pearson相关系数定义,AC与CNL具有强相关性,对模型最终结果有着几乎相同的影响,因此,舍弃CNL。再结合XGBoost特征重要性排序结果,CAL、SP和GR得分较靠前,也属于重要参数。因此将AC(声波时差)、CAL(井径)、DEN(密度)、GR(自然伽马)、SP(自然电位)作为神经网络的输入参数,这样可以避免规模过大的问题,保证更快的运行速度和更高的精度。

表1 特征重要性排序

测井曲线	重要性得分
AC	1 004.0
CAL	838.0
SP	786.0
GR	697.0
CNL	667.0
DEN	532.0

### 3.4 数据归一化

本研究的主要目的是利用常规测井资料对储层脆性指数进行预测。因此,将测井曲线作为网络的输入,而将脆性值作为网络的输出。考虑到每一种测井曲线的量纲不一,所以要进行无量纲化处理,通过归一化转换将数据集中的数据缩放到一个区间。采用常用的min-max归一化方法,将数据线性缩放到0~1之间的值:

$$\bar{x}_i = \frac{x_i(k) - \min\{x_i(k)\}}{\max\{x_i(k)\} - \min\{x_i(k)\}}, i = 0, 1, \dots, m \quad (13)$$

其中,  $\bar{x}_i$  为标准化处理结果,  $x_i(k)$  表示原始数据,  $\max\{x_i(k)\}$  和  $\min\{x_i(k)\}$  分别表示原始数据集的最大值和最小值,归一化后数据集部分数据如表2所示。

表2 数据集部分数据

AC	CAL	CNL	DEN	GR	SP	BI
0.50	0.75	0.29	0.21	0.45	0.92	0.38
0.00	0.83	0.57	0.47	0.31	0.91	0.37
.....	.....	.....	.....	.....	.....	.....
0.38	0.78	0.31	0.22	0.52	0.89	0.36
0.57	0.16	0.38	0.49	0.36	0.91	0.40

之后是数据集的划分,将数据集以7:3的比例划分为训练集和测试集。因此在这个模拟中,大约70%的数据用于训练,30%用于测试。为了评价预测结果,以均方误差(MSE)和均方根误差(RMSE)作为性能

指标。表达式如下:

$$MSE = \frac{1}{n} \sum_{i=1}^n (f(x_i) - y_i)^2 \quad (14)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f(x_i) - y_i)^2} \quad (15)$$

其中,  $y_i$  为测量数据,  $f(x_i)$  为预测数据,  $n$  为用于训练或测试神经网络的样本个数。

### 3.5 模型参数选择与训练

为了保证实验公平,对涉及到的神经网络模型统一采用 Adam 优化器优化内部参数。Adam 是一种基于低阶矩自适应估计的随机目标函数一阶梯度优化算法,该算法易于实现,计算效率高,内存需求小。Adam 通过优化目标函数,也适用于 BiLSTM 神经网络,其目标是找到一组使均方误差最小化的参数<sup>[15]</sup>。

在 CNN-LSTM-AM 模型实验过程中,需要对神经网络的超参数进行调整,以保证神经网络的性能尽可能最优,影响模型效果的超参数主要有四个:批处理参数、训练步数、学习率以及激活函数。超参数优化最常用的是人工搜索,人工搜索需要耗费较大的时间和精力。若通过正交试验法获得最优组合超参数,将耗费大量的资源,并且不能确保得到的超参数组合是最优解。因此,该文选取经典群智能算法-粒子群算法进行超参数优化。

表 3 神经网络超参数选择情况

超参数	实验设置
批处理参数	{ 16-128 }
训练步数	{ 10-90 }
学习率	{ $10^{-4}$ - $10^{-1}$ }
隐层神经元个数	{ 32-256 }

经粒子群寻优后,当批处理参数为 32,训练步数为 50,学习率为 0.005,隐层神经元个数为 64 时,模型效果最好。

### 3.6 实验结果与分析

为了进一步验证文中模型在脆性预测中的效果,选取集中传统机器学习方法进行对比,并与文献[9]提出的方法进行了对比。同时为了减少实验误差,将模型运行 10 次,取平均值为最终预测结果,因测试集数量较多,不好进行观察,所以选择了部分测试集的预测进行结果展示,预测结果对比如图 7~10 所示,横坐标是测试集的数量,纵坐标是脆性指数。

总体来说,CNN-BiLSTM-AM 模型得到的低偏差说明了其具有较好的预测能力。虽然传统的机器学习算法在预测过程中不如 CNN-BiLSTM-AM 强,但是它们仍然可以被认为是一种较为可靠的脆性预测方法。

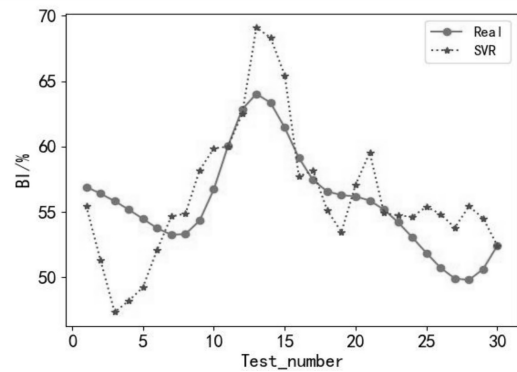


图 7 SVR 预测结果

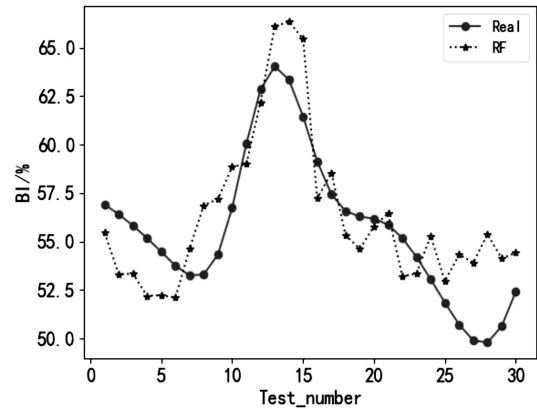


图 8 随机森林预测结果

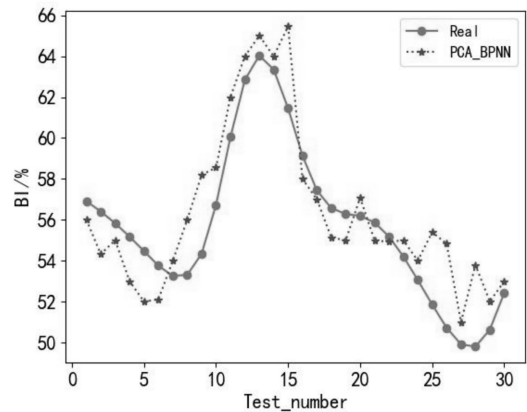


图 9 PCA-BPNN 预测结果

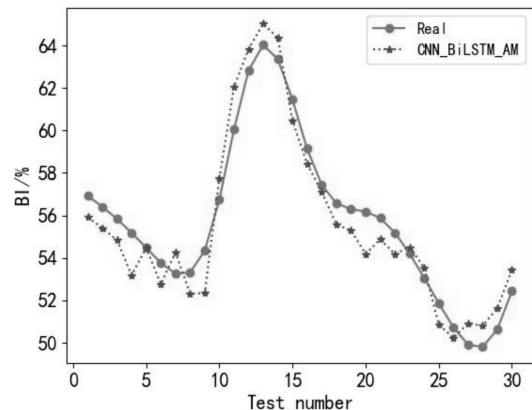


图 10 CNN-BiLSTM-AM 模型预测结果  
在训练和测试过程中,CNN-BiLSTM-AM 在

MSE 和 RMSE 误差方面(见表 4)都比传统机器学习模型表现更好。实验结果表明,该模型在预测岩石脆性精度方面要更加优于传统机器学习模型,符合实验的预期成果。传统方法忽略了测井曲线内部信息的关联,BiLSTM 可以有效地考虑上下文信息,因此会比传统的方法更为精确,另外,在双向长短期的基础上添加了一维卷积进行特征提取,同时也添加了注意力机制,可以有针对性地对特征分配权重,从而有更小的误差。

表 4 各模型的 MSE 和 RMSE

模型	MSE	RMSE
支持向量回归	0.007 96	0.088 88
随机森林	0.006 21	0.078 80
PCA-BPNN	0.004 96	0.070 42
CNN-BiLSTM-AM	0.003 50	0.059 16

#### 4 结束语

该文讨论了传统的机器学习模型和 CNN-BiLSTM-AM 模型在预测页岩储层脆性方面的应用,得出了以下结论:

(1)CNN-BiLSTM-AM 模型选取多条常规测井曲线作为输入变量,脆性指数作为输出变量,成功地对岩石脆性进行了准确、可靠的脆性指数预测,为储层脆性预测提供一种新途径。

(2)传统的机器学习模型和 CNN-BiLSTM-AM 模型都可以进行脆性预测,但是 CNN-BiLSTM-AM 模型的误差小于其他模型,并且该文运用相关系数矩阵分析等数据挖掘技术有助于提取特定数据集的最重要和最相关的特征。

(3)建立的 CNN-BiLSTM-AM 模型较为复杂,所以运行速度比其他模型慢,仍可以进一步对其进行优化,提升预测的速度和效率。

#### 参考文献:

- [1] WANG Y, WATSON R, ROSTAMI J, et al. Study of borehole stability of Marcellus shale wells in longwall mining areas[J]. J Pet Explor Prod Technol, 2014, 4: 59-71.
- [2] SHI X, LIU G, CHENG Y, et al. Brittleness index prediction in shale gas reservoirs based on efficient network models[J]. Journal of Natural Gas Science and Engineering, 2016, 35: 673-685.
- [3] 叶亚培,唐书恒,郝兆栋,等.一种基于测井数据的页岩脆性指数预测方法:中国,202110195831.9[P]. 2021-02-22.
- [4] HEIDARI M, KHANLARI G R, TORABI-KAVEH M, et al. Effect of porosity on rock brittleness[J]. Rock Mech Rock Eng, 2014, 47: 785-790.
- [5] JIN Xiaochun, SHAH S, TRUAX J. A practical petrophysical approach for brittleness prediction from porosity and sonic logging in shale reservoirs[C]//Society of petroleum engineers annual technical conference and exhibition. Amsterdam: [s. n.], 2014: 1-18.
- [6] LAI J, WANG G, HUANG L, et al. Brittleness index estimation in a tight shaly sandstone reservoir using well logs[J]. Journal of Natural Gas Science and Engineering, 2015, 27: 1536-1545.
- [7] SHI Xian, LIU Gang, JIANG Shu, et al. Brittleness index prediction from conventional well logs in unconventional reservoirs using artificial intelligence[C]//International petroleum technology conference. Bangkok: [s. n.], 2016.
- [8] 袁思乔,刘之的,王鼎涵.利用多测井参数预测致密砂岩脆性指数方法研究[J].地球物理学进展,2022,37(1):312-319.
- [9] YE Yapei, TANG Shuheng, XI Zhaodong. A new method to predict brittleness index for shale gas reservoirs; insights from well logging data[J]. Journal of Petroleum Science & Engineering, 2022, 208: 109431.
- [10] SUN D, LONBANI M, ASKARIAN B, et al. Investigating the applications of machine learning techniques to predict the rock brittleness index[J]. Applied Sciences, 2020, 10(5): 1691.
- [11] HAO Y, GAO Q. Predicting the trend of stock market index using the hybrid neural network based on multiple time scale feature learning[J]. Applied Sciences, 2020, 10(11): 3961.
- [12] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Trans Neural Network, 2002, 5(2): 157-166.
- [13] 武中原,张欣,张春雷,等.基于 LSTM 循环神经网络的岩性识别方法[J].岩性油气藏,2021,33(3):120-128.
- [14] 彭玉青,乔颖,陶慧芳,等.融入注意力机制的 PM2.5 预测模型[J].传感器与微系统,2020,39(7):44-47.
- [15] 林旭梅,胡川,朱广辉,等.基于深度残差反向传播神经网络的钢筋腐蚀检测[J].科学技术与工程,2021,21(33): 14351-14355.