

结合地质导向作业的实时智能地层识别预测模型

陈嘉豪¹, 李谦^{*1,2}, 曾小龙¹, 罗浩天¹, 邹欣悦¹

(1. 成都理工大学环境与土木工程学院, 四川 成都 610059;
2. 地质灾害防治与地质环境保护国家重点实验室(成都理工大学), 四川 成都 610059)

摘要: 地质导向是当前钻探智能化的核心组成部分, 但当前地质导向中关键的地层识别还缺乏算法模型支撑。本文采用深度学习算法, 基于四川威远区块实际采集的井身轨迹参数数据、测井数据, 建立了能够实时预测钻遇地层的算法体系。该体系首先通过时序预测算法完成对后续井身轨迹的预测, 其次基于该预测值实现对地层伽马值的预测, 最后根据地层伽马值完成对钻遇地层的类型识别。为提高整个算法体系预测准确性, 各环节采用多种算法进行对比, 结果证明采用循环神经网络训练的实时预测, 随机森林预测伽马值回归, 支持向量机预测地层分类的训练效果最好, 预测和识别精度分别达到了 0.81、0.83、81.8%。地质导向过程中对实时钻遇地层的预判, 为现场提供数据支撑以及辅助判断的手段, 帮助实际工程更可靠有效地实现地质目标。

关键词: 地层预测识别; 地质导向; 实时预测; 机器学习

中图分类号: P634; TE24 **文献标识码:** A **文章编号:** 2096-9686(2023)S1-0135-08

Integration of real-time intelligent stratigraphic identification and prediction model for geologically-guided field operation

CHEN Jiahao¹, LI Qian^{*1,2}, ZENG Xiaolong¹, LUO Haotian¹, ZOU Xinyue¹

(1. College of Environment and Civil Engineering, Chengdu University of Technology,
Chengdu Sichuan 610059, China;

2. State Key Laboratory of Geohazard Prevention and Geoenvironment Protection
(Chengdu University of Technology), Chengdu Sichuan 610059, China)

Abstract: This paper employs deep learning algorithms based on actual wellbore trajectory parameter data and well logging data collected from the Weiyuan block in Sichuan. An algorithmic framework capable of real-time predicting encountered strata during drilling is established. Firstly, this framework utilizes a time-series prediction algorithm to forecast the subsequent wellbore trajectory. Subsequently, based on this prediction, it achieves the forecast of strata gamma values. Finally, the identification of encountered strata types is accomplished by utilizing the predicted strata gamma values. To enhance the predictive accuracy of the entire algorithmic framework, various algorithms are employed and compared at each stage. The results demonstrate that real-time prediction using recurrent neural networks, gamma value regression prediction through random forests, and strata classification prediction using support vector machines yield the best training outcomes. The predictive accuracies reach 0.81, 0.83 and 81.8%, respectively. The real-time anticipation of encountered strata during the geologically guided process provides on-site data support and auxiliary judgment means. This contributes to the more reliable and effective achievement of geological objectives in

收稿日期: 2023-05-30; 修回日期: 2023-08-09 DOI: 10.12143/j.ztgc.2023.S1.020

基金项目: 川庆钻探钻采工程技术研究院项目“地质导向辅助决策技术研究”(编号: CQCDLG-2020-06)

第一作者: 陈嘉豪, 男, 汉族, 1999 年生, 硕士研究生, 地质资源与地质工程专业, 研究方向为人工智能在钻井中的应用, 四川省成都市二仙桥东三路 1 号, 940321650@qq.com。

通信作者: 李谦, 男, 汉族, 1987 年生, 副教授, 地质工程专业, 博士, 长期从事数据挖掘与人工智能在钻井中的应用研究工作, 四川省成都市二仙桥东三路 1 号, liqian2014@cdut.edu.cn。

引用格式: 陈嘉豪, 李谦, 曾小龙, 等. 结合地质导向作业的实时智能地层识别预测模型[J]. 钻探工程, 2023, 50(S1): 135-142.

CHEN Jiahao, LI Qian, ZENG Xiaolong, et al. Integration of real-time intelligent stratigraphic identification and prediction model for geologically-guided field operation[J]. Drilling Engineering, 2023, 50(S1): 135-142.

practical engineering.

Key words: stratigraphic prediction and identification; geologically-guided; real-time prediction; machine learning

0 引言

在钻井作业中,地质导向通过控制井身轨迹,以达到实际作业区域并满足实际工程指标^[1]。地质导向技术离不开测井技术与导向技术,其中常用的测井技术方法,有电阻率测井与伽马射线测井。其中伽马射线测井技术的优点更明显,该技术探测地层情况的边界信息更灵敏,分辨率清晰,常运用在一些非常规井导向中,且效果显著^[2-4]。导向技术方面,借助于随钻测井(LWD)导向技术,能够更迅速的实现钻探作业,并根据LWD测量的信息变化,维持较高的工作效率,并快速达到靶区^[5-6]。目前,我国处于自动化向智能化发展的阶段,传统的地质导向,不可避免的存在局限性。比如传统地质导向技术在垂直井与水平井的表现不一,LWD技术在水平井径向测量无法区分上下边界。进行新型智能导向的研究具有广阔的应用前景与研究价值。

许多专家学者已经在智能化导向钻井方面取得了不少成果。Qin等^[7]提出了一种多探头伽马射线测井方法,旨在通过对边界距离的迅速预测,以辅助导向工作的进行。Lu等^[8]提出了一种新的并行多链方法来进行反演,降低了传统的反演方法的成本。Kullawan等^[9]提出了一种离散随机动态规划模型,用以确定优化地质导向操作中的顺序决策,在实际中取得了较好的应用。Zong等^[10]提出了一种伽马射线LWD三维快进方法,可以实现快速方位成像,且泛用性较广。

地质导向需要从实时数据中推断,然而随钻数据却提供滞后的储层信息,并运用于之后的地质导向决策判断^[11]。因此,为了进一步辅助钻井作业中的导向工作,决定采用机器学习与实际数据结合训练。训练出的模型能在原始数据上直接预测未来随钻数据的数值。根据模型训练的结果,修正提高精度,可以在一定程度上实时预测井身的发展轨迹。本文提出了一种基于时序网络预测井斜发展轨迹模型,再根据井斜参数预测伽马值,最后根据伽马值预测地层分类。3个模型串联预测,最大程度提供实时数据,以支撑地质导向决策。

1 原始数据集

本文所用的原始数据集采集于威远地区页岩气田6口井。具体的6口井井号,以及采集数据如表1所示。

表1 采集数据参数

井号	Z201、Z202、Z203、Z204、W202、W204
井斜轨迹参数	MD(斜深)、TVD(垂深)、INCL(井斜角)、AZ-IM(方位角)、X(大地坐标)、Y(大地坐标)
测井参数	GR(伽马值)、DEP(深度)、地层分层

该地区的产层主要为五峰—龙马溪组。五峰组顶界主要为观音桥段介壳灰岩,厚度在1 m以内,之后为五峰组碳质硅质页岩。龙马溪组自上而下划分为龙二段与龙一段。龙一段又划分为龙一₂亚段与龙一₁亚段。龙一₁亚段特征集中表现为黑色碳质、硅质的页岩。当前钻井区域目的层为龙一₁亚段中的4个小层。由新到旧依次为龙一₁⁴、龙一₁³、龙一₁²、龙一₁¹。地层情况如图1所示。



图1 地层分布

2 算法介绍

2.1 建模流程

具体建模流程如图2所示,首先建立根据实时井身轨迹参数预测未来轨迹参数的模型,接着建立

以井身轨迹参数输入、伽马数据输出的伽马值预测模型,最后建立以伽马数据输入、地层分类参数输出的地层分类预测模型。旨在通过第一个模型中得到的未来预测井身轨迹参数数据,加入到伽马预测模

型、地层分类预测模型中,以得到实时预测的伽马值与地层分类值。这3种模型的建立分别需要深度学习中的回归器、分类器实现。

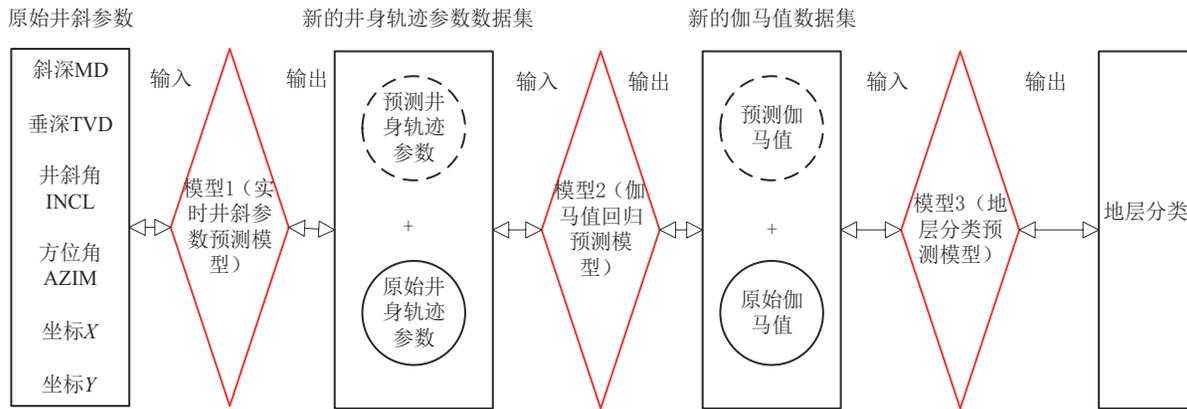


图2 模型工作示意

2.2 常见算法的用途

常见算法中的学习器包括分类器与回归器,回归器在连续样本值点建模的应用广泛,与分类器不同,回归器不需要提前进行人为分类,其模型的输入输出是具体的数值,而非标签类别。本文拟建模的伽马值预测就需要回归器得以实现。与地层类别参数不同,实际采得的伽马值为具体的数值,无法进行标签一一对应。采用回归器对伽马值进行预测回归,旨在通过对误差的衡量来寻求伽马值与其他输入参数变量之间的一种关系。通过回归器训练,可以输出预测的伽马值。

回归器常使用回归算法建模,评价模型的指标通常使用均方根误差(MSE)或 R^2 。其中MSE越接近于0, R^2 越接近1,模型的效果越好。具体公式如下:

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

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

式中: y_i ——观测值; \bar{y} ——观测值的平均值; \hat{y} ——回归模型第*i*个的观测值; n ——样本数量。

分类器的职责旨在根据输入数据进行标签化划分,同时输出结果为类别的标签。这也就意味着在输入数据时,需要人为根据数据特点进行分类,并给

与对应的标签类别。本文拟建模的地层分类预测模型,很好的对应分类器的特点。将实际采集到的地层具体分类信息,打上与之对应的数字标签,输入到分类器算法中进行学习建模,分类器可以在学习的过程中对地层分类的标签进行识别训练并输出预测结果。

分类器通常使用的是分类算法,评价分类器的模型常使用准确率(Accuracy)来表示,具体公式如下:

$$Accuracy = \frac{TP}{N} \quad (3)$$

式中:Accuracy——准确率;TP——分类结果中正确样本的数量; N ——所有的样本。

2.3 常见算法的类型

早先已有数据科学家明确指出,不存在一种万能算法适用于解决所有难题。因此,在实际解决问题时,需要根据实际问题类型,变量的数量,来选取合适的算法以解决问题^[12]。尽管这些特定的模型可能仅在数据采集的地区内具有一定的效用,但本文提出的建模流程思路可以在一定程度上帮助解决实际问题。本文将主要使用支持向量机、神经网络和随机森林等算法进行对比建模。

支持向量机(SVM)是一种简单、高效的监督算法,可用于解决分类与回归建模^[13]。SVM在回归方面更具特点,可以执行线性或非线性回归,并且在中小型、非线性数据集中表现优异^[14]。该算法通过向

量点判断是否分隔,向量点离该平面越远,正确分隔点的概率也就越大^[15-16]。此外,SVM还可以加入优化算法进一步优化模型,比如常见的灰狼算法、蚱蜢优化算法^[17-18]。

神经网络是深度学习中常见的算法,其功能强大,在多个领域得到广泛运用,比如信号处理、特征保留等^[19]。神经网络具有多种网络结构,常见的包括人工神经网络(ANN)、循环神经网络(RNN)、卷积神经网络(CNN)等^[20]。其工作原理是通过构建包含输入层、多层处理层、输出层组成的计算模型,同时调整层之间的权重,以得到最好的结果^[21]。神经网络不仅可以运用于分类任务、回归任务,在结合时间序列预测上也同样表现不俗,其中的RNN循环神经网络就是运用广泛的时序预测算法。

随机森林(RF)是目前运用最广泛的学习器之一^[22],其特点在于结合若干个弱学习器(决策树)组成一个强大的学习器,用于最终投票得分后的结果并输出^[23-24]。其集成弱学习器的处理方式在很大程度上是不稳定的,更依赖于数据集本身的结构特点,数据集自身微小的变化可能在决策上导致决策变化差异性明显^[25]。在最终输出结果上,通常随机森林表现出的结果较好。其原因在于集成的强学习器可以使得模型的结果更有效,同时克服单一学习器的局限,泛化性更广^[26]。

这3种常见的机器学习算法具备同时处理回归预测与分类预测的能力,但仍有一些常用的算法在单一的分类与回归预测中表现不俗。比如梯度回归算法、决策树分类算法、KNN分类算法。

决策树分类是单一的分类易受噪音影响适用于较小的数据集,决策树分类的核心思想是在数据特征空间中创建一棵树,其中每个内部节点表示一个特征条件,每个叶子节点代表一个类别。而随机森林分类是集成分类器,集成多个决策树进行分类。用单一的决策树分类根据结果可以预先了解模型的准确性下限,以及判断数据集中数据量的量级。为后面集成学习提高模型精确度有一定辅助作用。

梯度回归算法是一种迭代优化算法,通过迭代优化过程,逐步调整模型参数,以使损失函数逐渐减小。该过程通过计算损失函数关于模型参数的梯度,朝着梯度下降进行参数优化。使预测结果更接近实际观测值。

KNN分类是一种有监督分类算法,通过给定一

个未标记的数据点,根据其已与已知标记数据点之间的距离,将其分类为与其最近的 K 个已知标记数据点中占比最多的类别。KNN分类的关键在于 K 值的选取。较小较大的 K 值都会带来决策的边界模糊,通过引入轮廓系数对比,可以在实际运用中选取合适的 K 值。

在本文建模过程中3类算法的选择具体如表2所示。

表2 建模使用的具体算法

建模模型	算 法
时序预测模型	RNN神经网络
	随机森林回归
	梯度回归
伽马值回归预测模型	支持向量机回归
	ANN神经网络
	决策树分类
地层分类预测模型	随机森林分类
	支持向量机分类
	KNN分类

3 模型建立

3.1 实时预测井身轨迹参数模型

在地质导向过程中,实时预测对于保持储层最大接触率、控制井眼轨迹、确保储层遇钻率等有着重要的意义。建立时间序列与预测参数对应的模型,将井身轨迹参数作为输入,结合特定的时序预测算法,可输出同时间点未来若干时间步数的井身轨迹参数。通常,模型的准确率与数据集自身特点有关,意味着数据集的样本点越多,预测的效果越好。随着钻进的进行,新采集的数据可以不断加入已经训练好的数据,并进一步用于预测该时间点后的未来时间步的轨迹参数。模型的精确度会越来越高。具体建立井身轨迹参数预测模型如图3所示。

在时序预测中,循环神经网络(RNN)是一种十分常见的学习算法。如图4示,随着神经网络训练中超参数迭代次数的增加,模型的MSE值也会趋于稳定。其中,当迭代到某些次数时,MSE值表现出了一些升高又再次降低,这是因为在训练过程中学习率变化与局部优化点所导致的,并不能反应模型的效果变差,反应的是该模型的动态训练过程。图4中MSE值迭代次数在200次左右已经接近

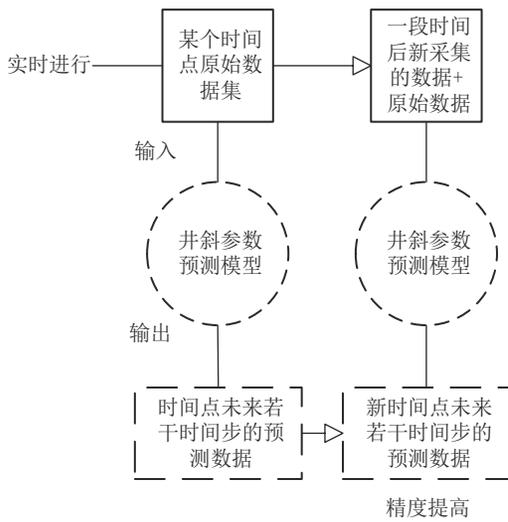


图 3 井身轨迹参数预测模型工作示意

0.00025,模型的精准度已较好的达到了要求。

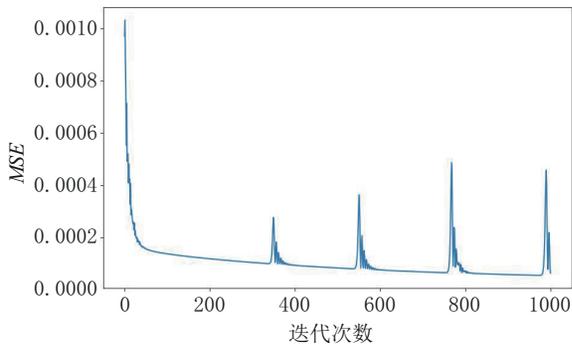
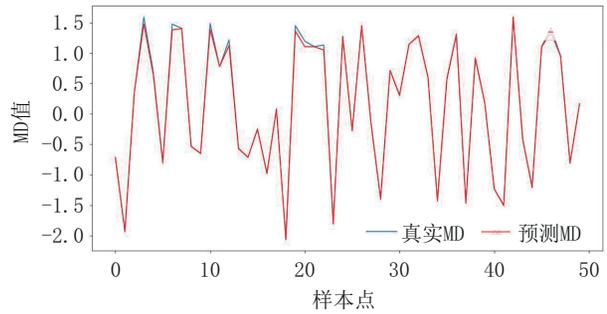


图 4 井斜参数预测模型的 MSE 的迭代次数

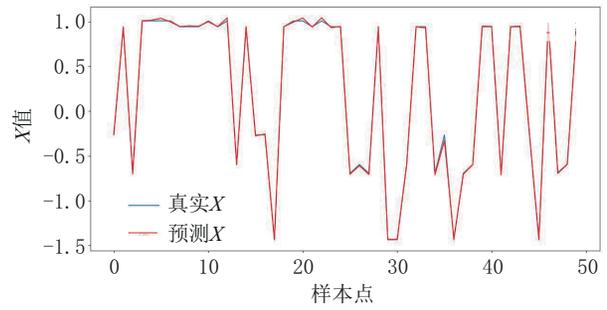
评价模型的好坏不能单一参考 MSE 值,也要根据实际与预测的对比。从测试集上得到训练结果 R^2 达到 0.81,在测试集中随机选取 50 个样本点进行训练,得到效果如图 5 所示,用于预测的多个参数表现出了较好的重合,且参数的预测走势一致。可以在一定程度上说明该模型的效果较好。

3.2 伽马值预测模型

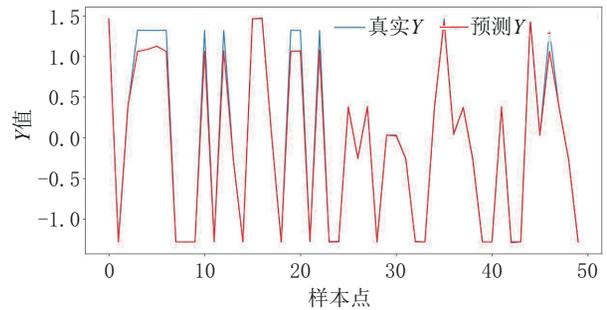
本文已经建立了时间序列与未来井身轨迹参数的预测模型,而对未来深度对应的伽马值进行预测,属于回归预测。而井身轨迹参数、深度、伽马值是互相对应的,具体如图 6 所示。将井身轨迹参数作为输入,而非深度参数输入,伽马值参数作为输出。这样的好处有 2 点,其一是井身轨迹参数的未来值可以直接通过模型 1 得到,便于实时预测伽马值;其二是井身轨迹参数作为输入,其特征向量有 6 个,若直接使用深度输入特征向量只有 1 个。用于输入的特



(a)MD 对比



(b)X 值对比



(c)Y 值对比

图 5 测试集随机选取真实值预测值对比

征向量参数之间彼此越不相关,增加特征向量个数,其得到的模型精确度越高。输入参数之间的相关性值如表 3 所示,可以看到相关性矩阵反映出参数之间的相关性较低。可以考虑用输入参数代替深度作为输入。

本文将井身轨迹参数作为输入,伽马值作为预测变量进行建模。如图 7 所示,采取多种算法所得到的结果显示,RFR 与 GBR 模型的评价较高,随机森林的 R^2 达到了 0.83,梯度回归则达到了 0.79。而支持向量机与人工神经网络的模型效果相对较差。

模型训练效果如图 8 所示,对角线代表理论线,若散点落聚在理论线上,则可以说明该模型的效果较好。在测试集上绘制出的训练效果中,其中

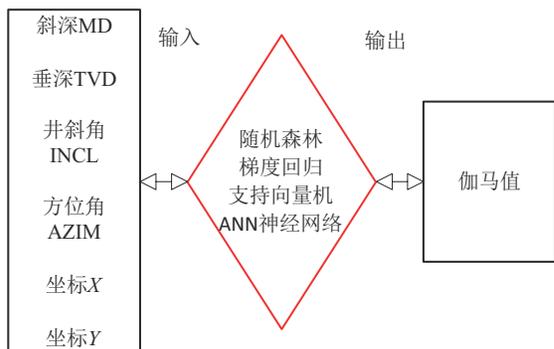


图6 伽马值模型输入输出参数

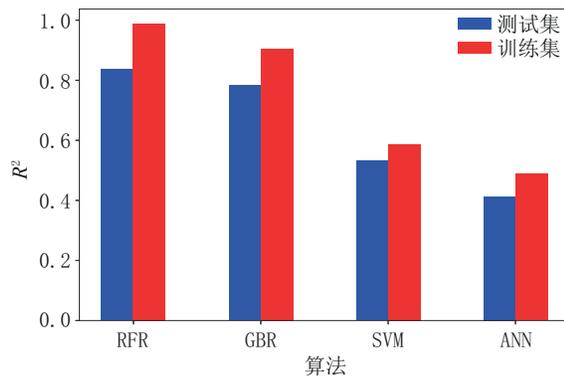


图7 4不同算法伽马值回归模型R²

ANN、SVM分布在直线两侧,RF与GR所建立的模型聚落在理论线上,RF在测试集上的R²达到了0.676,结合两图说明RF建立的回归预测模型精度更高。

3.3 地层分类预测模型

最后,根据伽马值与其对应的地层分类,将伽马

值作为输入参数,地层分类作为输出建立模型,将地层分类情况打上标签。

如图9所示,采取决策树(DT),随机森林(RF),支持向量机(SVM),K近邻(KNN)分类器得到的模型准确率在75%左右。其中RF与SVM的

表3 输入参数之间相关性矩阵

参数	MD	X	Y	TVD	AZIM	INCL
MD	1	0.039424	0.171330	0.999999	0.281378	0.319773
X	0.039424	1	-0.147518	0.039713	0.350433	-0.292032
Y	0.171330	-0.147518	1	0.171347	0.156450	0.045038
TVD	0.999999	0.039713	0.171347	1	0.281495	0.318880
AZIM	0.281378	0.350433	0.281495	0.281495	1	0.121834
INCL	0.319773	-0.292032	0.318880	0.318880	0.121834	1

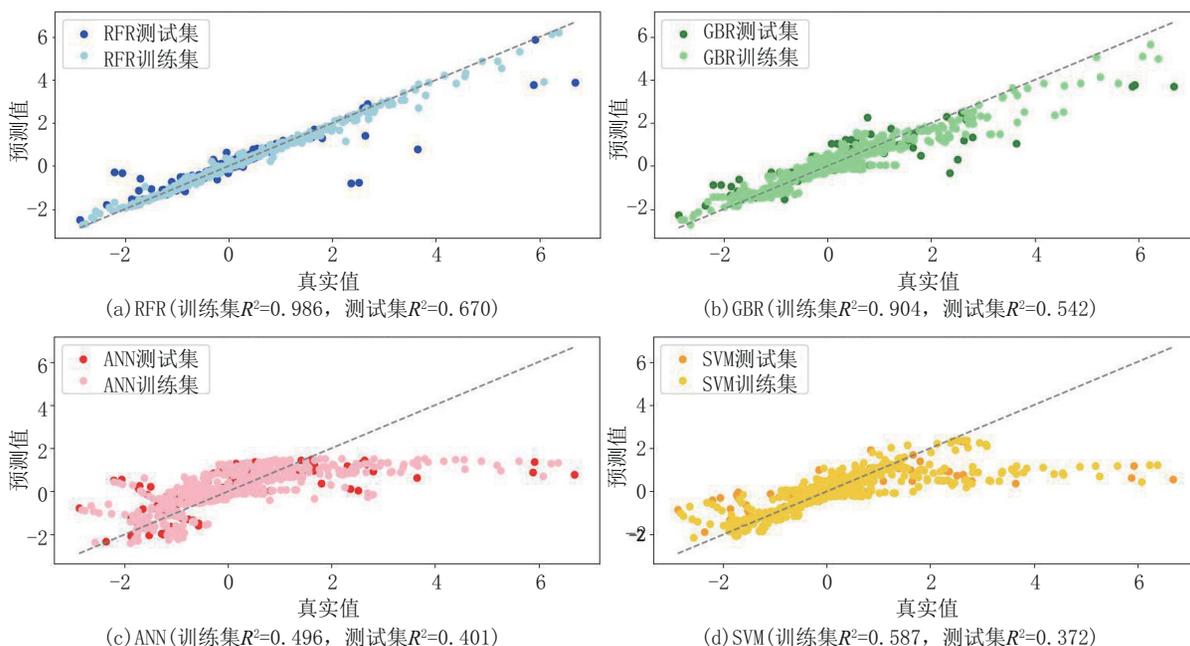


图8 4不同算法模型训练效果

准确率达到了81.8%。表现出了较好的精度。

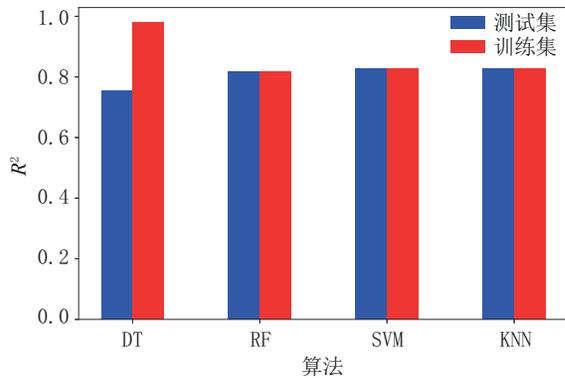


图9 4分类器训练模型准确率

图9的结果中RF与SVM的模型在整体准确率上差距不大。基于数据集中的特点,分类器在龙一层中的表现也是重点关注对象。引入多分类的混淆矩阵以配合说明,具体如表4所示。其中多分类的混淆矩阵是一种常见的具体化的评价多分类模型精确度的指标,可以具体反应每一个标签的预测结果。该矩阵的每一行代表一个标签类。以支持向量机为第一列为例子,其中第一行第一列从左往右依次为:将42个样本分类为标签五峰组(0),1个样本分类为龙一¹(1),0个样本分类为龙一²(2),0个样本分类为龙一³(3),1个样本分类为龙一⁴(4),其中准确率为95%。SVM在龙一¹(1)表现出色,准确率高达90%,在龙一⁴(4)高达99%。但在龙一²

(2)与龙一³(3)中都将样本更多的分类到了龙一⁴(4)中。

3.4 模型总结

威远地区辅助地质导向建模推荐流程如下:

首先借助RNN神经网络训练并建立实时井身轨迹参数预测模型,输入实时的井身轨迹参数,得到未来井身轨迹参数,该模型的精确度MSE指标可达到0.00025, R^2 达到0.81精度较高。

然后运用随机森林回归器训练并建立伽马值预测模型,该模型输入井斜参数数据,可得到预测的伽马值,随机森林模型的 R^2 在训练集达到了0.986,测试集达到了0.67,模型的精度尚可。

最后运用支持向量机分类器建立地层预测参数模型,输入伽马值得到预测的地层分类。

该模型的准确度达到了81.8%,在目的层龙一¹与龙一⁴中分类表现出色。具体如图10所示。

4 结论

(1)基于威远地区钻井的数据,实现了基于时序的井身轨迹参数预测、测井伽马值预测与地层预测的模型建立。确定了从井身轨迹—伽马值—地层分类,依次对应的预测模型建立流程。

(2)分别选取了各自模型结果最好的算法,分别是采用RNN循环神经网络的井身轨迹参数预测模型, R^2 达到0.81。采用随机森林的伽马值预测模型, R^2 达到0.83。以及采用支持向量机的地层类别

表4 RF与SVM的混淆矩阵

分类标签	随机森林					支持向量机				
	0	1	2	3	4	0	1	2	3	4
五峰组(0)	43	0	0	0	1	42	1	0	0	1
龙一 ¹ (1)	0	14	4	1	2	0	19	1	0	1
龙一 ² (2)	0	3	4	6	20	0	2	5	10	16
龙一 ³ (3)	0	5	3	4	19	0	2	4	9	16
龙一 ⁴ (4)	0	0	0	1	158	1	0	0	0	158

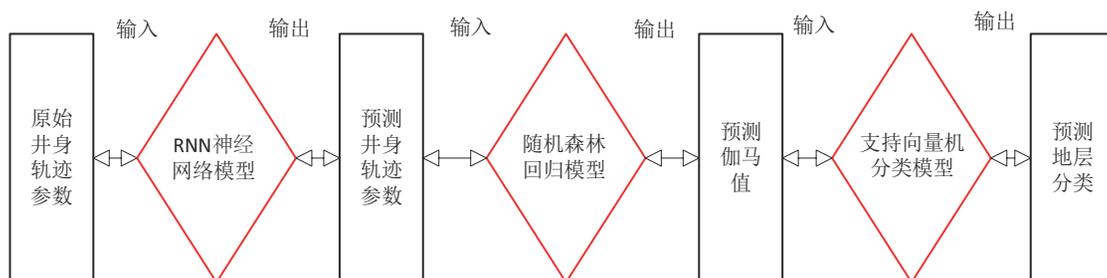


图10 建模推荐使用流程

预测模型,准确率达到81.8%,在钻采目的层龙一¹与龙一⁴中表现良好。

参考文献:

- [1] Jin Y, Wan Q, Wu X, et al. FPGA-accelerated deep neural network for real-time inversion of geosteering data[J]. *Geoenergy Science and Engineering*, 2023,224:211610.
- [2] Mohamed D, Jeremy C. Unconventional horizontal well placement challenges—LWD geosteering applications utilizing a unique spectral azimuthal gamma ray in real time[Z]. 2014.
- [3] Thwaites N, Suh A. Use of near bit azimuthal gamma ray and inclination tool improves geosteering in CBM wells, Airth Field, Scotland[Z]. EAGE Publications BV, 2014:1-12.
- [4] Qin Z, Pan H, Wang Z, et al. A fast forward algorithm for real-time geosteering of azimuthal gamma-ray logging [J]. *Applied Radiation and Isotopes*, 2017,123:114-120.
- [5] Moorkamp M, Lelièvre P G, Linde N, et al. *Integrated Imaging of the Earth: Theory and Applications*[M]. John Wiley & Sons, 2016.
- [6] Sun Q, Duan Y, Yan Y, et al. Forward Algorithm of LWD Azimuthal Gamma [J]. *Journal of System Simulation*, 2018, 30(11):4268.
- [7] Qin Z, Tang B, Wu D, et al. A qualitative characteristic scheme and a fast distance prediction method of multi-probe azimuthal gamma-ray logging in geosteering [J]. *Journal of Petroleum Science and Engineering*, 2021,199:108244.
- [8] Lu H, Shen Q, Chen J, et al. Parallel multiple-chain DRAM MCMC for large-scale geosteering inversion and uncertainty quantification [J]. *Journal of Petroleum Science and Engineering*, 2019,174:189-200.
- [9] Kullawan K, Bratvold R B, Bickel J E. Sequential geosteering decisions for optimization of real-time well placement[J]. *Journal of Petroleum Science and Engineering*, 2018,165:90-104.
- [10] Zong C, Liu J, Liu Z, et al. A fast forward algorithm of azimuthal gamma imaging logging while drilling[J]. *Applied Radiation and Isotopes*, 2023,194:110659.
- [11] Dawoud A, El Mahdi A, Ayoub M, et al. Geosteering long horizontal holes in Abu Dhabi heterogeneous carbonate reservoirs[Z]. 2010128851.
- [12] Mahesh B. Machine learning algorithms—A review[J]. *International Journal of Science and Research (IJSR) (Internet)*, 2020,9:381-386.
- [13] 奉国和. SVM分类核函数及参数选择比较[J]. *计算机工程与应用*, 2011,47(3):123-124.
- [14] G E Ron A E L. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow* [M]. O’Reilly Media, Inc., 2022.
- [15] Altin F G, Budak O, özcan F. Predicting the amount of medical waste using kernel-based SVM and deep learning methods for a private hospital in Turkey[J]. *Sustainable Chemistry and Pharmacy*, 2023,33:101060.
- [16] Raghavendra. N S, Deka P C. Support vector machine applications in the field of hydrology: A review[J]. *Applied Soft Computing*, 2014,19:372-386.
- [17] Hamad A, Houssein E H, Hassanien A E, et al. A hybrid EEG signals classification approach based on grey wolf optimizer enhanced SVMs for epileptic detection[C]//Cham: Springer International Publishing, 2018.
- [18] Houssein E H, Hamad A, Hassanien A E, et al. Epileptic detection based on whale optimization enhanced support vector machine[J]. *Journal of Information and Optimization Sciences*, 2019,40(3):699-723.
- [19] Tang W, Gong Y, Qiu G. Feature preserving 3D mesh denoising with a Dense Local Graph Neural Network[J]. *Computer Vision and Image Understanding*, 2023,233:103710.
- [20] 胡中源,薛羽,查加杰. 演化循环神经网络研究综述[J]. *计算机科学*, 2023,50(3):254-265.
- [21] Qian K, Tian L. An embedded Hamiltonian dynamic evolutionary neural network model for high-dimensional data recognition [J]. *Applied Soft Computing*, 2023,141:110299.
- [22] Resende P A A, Drummond A E C. A survey of random forest based methods for intrusion detection systems[J]. *ACM Computing Surveys (CSUR)*, 2018,51(3):1-36.
- [23] Mohapatra N, Shreya K, Chinmay A. *Optimization of the Random Forest Algorithm*[M]. Springer, 2020:201-208.
- [24] 王奕森,夏树涛. 集成学习之随机森林算法综述[J]. *信息技术*, 2018,12(1):49-55.
- [25] Karabadi N E I, Amara Korba A, Assi A, et al. Accuracy and diversity-aware multi-objective approach for random forest construction [J]. *Expert Systems with Applications*, 2023, 225:120138.
- [26] 吕红燕,冯倩. 随机森林算法研究综述[J]. *河北省科学院学报*, 2019,36(3):37-41.

(编辑 周红军)