

遗传算法优化的智能分类器用于复杂碳酸盐岩岩性精细识别的对比

张鹏, 庄彬, 胡志昕

(成都理工大学能源学院, 四川 成都 610059)

摘要 针对川东开江-梁平海槽台缘带飞仙关组二段鲕粒滩相碳酸盐岩岩性复杂、变化不明确等问题,基于岩心、薄片、测井等资料,使用遗传算法优化的多种智能分类器,对比不同模型的精度与适用性,以探求精细识别研究区鲕粒滩相复杂碳酸盐岩岩性,解决研究区复杂岩性识别困难、精度不高的技术难题。研究表明:遗传算法优化的支持向量机分类器岩性识别结果客观可靠,识别准确率达到95.56%,优于其他智能分类器,与岩心、测井资料高度吻合,适用于碳酸盐岩复杂岩性识别,具有很好的推广应用前景。

关键词 测井数据;碳酸盐岩;复杂岩性识别;遗传算法;智能分类器

中图分类号:TP23

文献标识码:A

文章编号:1007-0745(2023)04-0028-03

位于川东北地区开江-梁平海槽台缘带飞仙关组二段碳酸盐岩岩性复杂,非均质较强,变化不明确,通过测井资料识别该段岩性研究较少。传统的交会图、矿物成分计算、判别分析等方法^[1-2]在该鲕粒滩相碳酸盐岩识别中效果并不突出,难以准确识别。本文综合利用元坝飞二段岩心、录井与测井资料,通过对比多种遗传算法(Genetic Algorithm, GA)优化的智能分类器模型的精度与适用性,选择分类效果最好的分类识别模型用于精细识别元坝飞二段鲕粒滩相复杂碳酸盐岩岩性,解决复杂碳酸盐岩岩性识别困难的技术难题。对比多种智能分类器效果发现,遗传算法优化的支持向量机智能分类器识别精度最高,整体识别准确率达到95.56%,可用于解决研究区测井岩性识别问题。

1 智能分类器方法

1.1 遗传优化算法

遗传算法由 John H. Holland 教授提出,通过模拟自然选择和遗传学机理的生物进化过程,搜索全局最优解的一种优化算法。遗传算法最大的特点就是将优化问题中的参数转换成了编码的个体,从而不需要考虑优化问题中的参数问题,只需要处理进行了编码的个体,采用概率化的寻优方法,不需要确定的规则就能自动获取和指导优化的搜索空间,自适应地调整搜索方向^[3-5]。

1.2 GA-BP 神经网络

BP (Back Propagation) 神经网络是 1986 年由 Rumelhart

和 McClelland 提出的一种按照误差逆向传播算法训练的多层前馈神经网络模型,具有高度非线性和较强容错能力,其优越的非线性映射能力和柔性的网络结构被广泛应用于分类问题,但其收敛速度较慢,易陷入局部极小值,初始的权值与阈值对分类器识别精度影响较大^[6]。所以针对传统 BP 神经网络存在的问题,基于遗传算法优势,通过遗传算法寻找全局最优的初始权值与阈值,优化 BP 神经网络参数,构建 GA-BP 神经网络智能分类器,以提高分类器识别准确度。

1.3 GA-SVM

支持向量机 (Support Vector Machine, SVM) 是建立在统计学习理论和结构风险最小化的基础上,具有坚实的理论基础和完善的理论体系,通过寻找最优超平面,实现精准分类并最大化超平面的分类间隔,在理论上充分保证了模型的泛化能力。传统支持向量机准确度的高低很大程度上取决于对核函数参数 g 和惩罚因子 c 的选择,参数选取不合理会直接影响预测精度^[7-10]。利用遗传算法全局寻优的优势,寻找最优的核函数值与惩罚因子值,构建 GA-SVM 智能分类器,以提高该智能分类器的识别精度,增强其岩性识别的泛化能力。本文采用具有较宽收敛域的径向基核函数,其公式 (1) 如下所示:

$$k(x_i, x_j) = e^{-g\|x_i - x_j\|} \quad (1)$$

其中, $k(x_i, x_j)$ 为核函数; g 为核参数,代表了

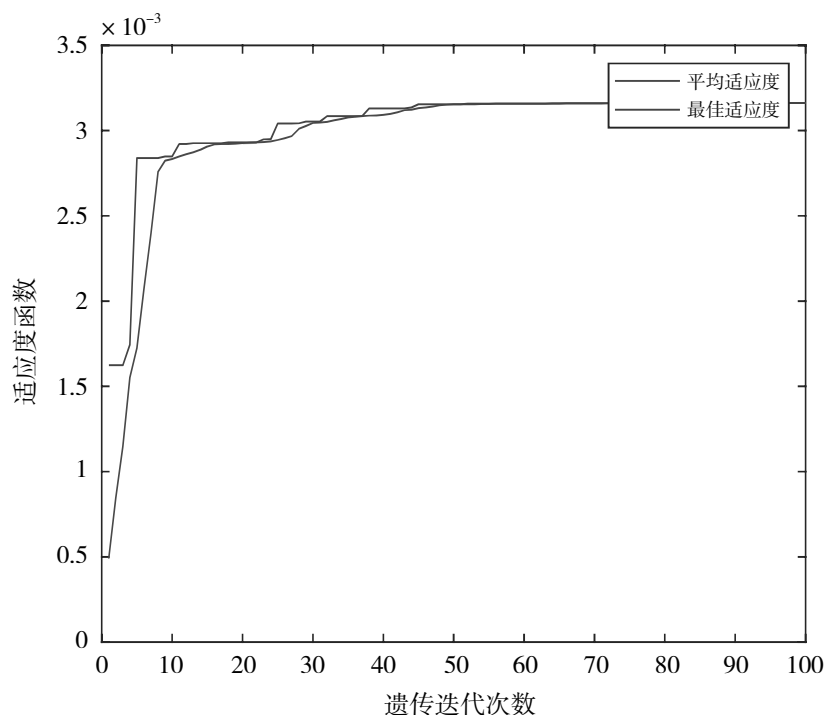


图 1 GA-BP 神经网络分类器遗传进化曲线图

径向基核函数作用的宽度。

2 岩性识别智能分类器的建立

2.1 数据及测井参数选择

元坝飞二段碳酸盐岩储层岩性较为复杂,以灰岩为主要发育岩性。基于录井和取心段资料,归纳出溶孔灰岩、鲕粒灰岩、微晶灰岩、砂屑灰岩、灰质白云岩、含泥灰岩和泥质灰岩共 7 种岩性,构建了 450 个可靠岩性数据样本,从中随机选择 390 个数据用于训练,剩余 60 个数据用以检验模型的适用性。不同测井曲线上任有一定的敏感性,可利用智能分类器进行岩性识别研究,所以选择自然伽马、深侧向电阻率、声波时差、中子和密度等 5 种测井参数作为输入特征值。

2.2 智能分类器训练

通过遗传算法寻优的初始参数统一设置为:进化代数为 100-300,种群规模为 20;交叉概率为 0.4,变异概率为 0.2。GA-BP 神经网络智能分类器初始设置网络训练最大迭代次数为 1000 次,训练目标精度预期为 0.02,学习率为 0.01,隐藏层节点数采用试错法确定。在 GA-SVM 智能分类器训练中,通过遗传算法寻优的最佳核函数值与惩罚因子值均通过 5 折交叉验证验证其适用性。

2.3 智能分类器训练结果分析

2.3.1 GA-BP 神经网络智能分类器结果分析

经过测试,传统的 BP 神经网络智能分类器与 GA-BP 神经网络智能分类器均在隐含层神经元节点数 15 时,达到目标精度要求。BP 神经网络智能分类器 MATLAB 计算精度 $r=0.90$;GA-BP 神经网络智能分类器 MATLAB 计算精度 $r=0.91$ 。GA-BP 神经网络智能分类器遗传进化至 60 代左右时最优,均方根误差最小,遗传进化曲线如图 1 所示。

传统 BP 神经网络智能分类器训练精度为 91.79%,检验精度为 85.00%;GA-BP 神经网络智能分类器训练精度为 92.56%,检验精度为 86.67%。GA-BP 神经网络智能分类器在研究区复杂碳酸盐岩岩性识别问题上精度更高。两种模型均在砂屑灰岩与微晶灰岩识别中判别错误较多,其原因可能为砂屑灰岩与微晶灰岩测井响应特征差异较小,导致两种智能分类器识别错误。此外,对于样本数量较少的灰质白云岩,GA-BP 神经网络智能分类器识别效果较好,仅判错 2 个样本点,正确率高达 90.0%,较传统的 BP 神经网络智能分类器 75.00% 的正确率明显提高。综上所述,GA-BP 神经网络智能分类器的岩性识别效果较好,有效避免陷入局部最小值的情况,提高了岩性识别精度。

2.3.2 GA-SVM 智能分类器结果分析

传统的 SVM 分类器训练精度 96.41%，检验精度为 86.67%；GA-SVM 分类器训练精度为 96.67%，检验精度为 88.33%，适应度函数值达到了 90.26%，遗传进化至 150 代左右时最优。两种智能分类器精度相差较小，GA-SVM 分类器训练精度和检验精度均略高于传统 SVM 分类器，但传统 SVM 分类器得到的最佳惩罚因子 c 为 168.89，最佳核常数 g 为 4.59；GA-SVM 分类器得到最佳惩罚因子 c 为 49.49，最佳核常数 g 为 9.32，两者相比，遗传算法优化后的分类器（GA-SVM），惩罚因子 c 明显减小，核参数值有所增大，分类器模型的复杂度降低，使得结构风险较小，增强了分类器的泛化能力，比传统的 SVM 分类器效果更好。

3 分类效果对比分析

为验证不同智能分类器在复杂碳酸盐岩岩性识别中的优势，在上述 GA-BP 神经网络智能分类器、GA-SVM 智能分类器的基础上，又采用了决策树法进行了岩性识别，并对其识别效果进行了对比分析。

3.1 决策树法

决策树是一种常见的智能分类器，是一种树状结构的分类模型^[11]。基于已建立的样本集，决策树法的训练精度为 77.69%，检验精度为 81.67%，总体精度较低，判错样本数量多。对砂屑灰岩和灰质白云岩样本岩性识别准确度仅为 37.21% 与 63.00%，难以达到精细识别要求，对复杂碳酸盐岩岩性识别问题适用性较差。

3.2 单井岩性识别效果分析

在相同样本训练的模型中，GA-BP 神经网络智能分类器总体准确率为 91.78%，GA-SVM 智能分类器准确率为 95.56%，GA-SVM 智能分类器判错数量明显减小，岩性识别准确率更高。对不同的岩性而言，GA-BP 神经网络智能分类器与 GA-SVM 智能分类器对溶孔灰岩、鲕粒灰岩和灰质白云岩识别准确度较高，对砂屑灰岩与微晶灰岩识别准确度均较低，但 GA-SVM 智能分类器识别效果更好，对砂屑灰岩识别精度提升至 74.42%，整体判错样本数更少，识别准确性较强，可以用于解决复杂碳酸盐岩岩性识别问题。

使用训练好的 GA-SVM 智能分类器，对取心井 Y21 取心段岩性识别，除在 6646.75-6648.40m 与 6649.44-6652.52m 段，GA-SVM 智能分类器将砂屑灰岩识别为微晶灰岩外，其余段均与取心段岩性一致；而且对样本数量较少的灰质白云岩段（6658.18-6660.13m）也能

准确识别；该智能分类器岩性划分更合理、更细致，应用效果较好。

4 结论

1. GA-BP 神经网络智能分类器有效避免模型的局部最优解，显著提升了模型对小样本岩性识别的准确率；GA-SVM 智能分类器优化了核常数与惩罚因子值，使得模型的结构风险较小，具有较强的泛化能力。

2. 常规的决策树法，对复杂碳酸盐岩岩性识别困难，不适用；GA-SVM 智能分类器训练精度与检验精度最高，整体准确率达到 95.56%，优于 GA-BP 神经网络智能分类器，对元坝飞仙关组二段复杂碳酸盐岩识别效果较好，具有很好的推广应用前景。

参考文献：

- [1] 叶榆,程超,蒋裕强,等.川东飞仙关组鲕粒滩岩性识别及其分布特征[J/OL].沉积学报:1-20[2023-04-06].
- [2] 蒋裕强,邓虹兵,易娟子,等.开江—梁平海槽西侧飞仙关组不同类型鲕滩储层特征及其控制因素研究[J].特种油气藏,2020,27(01):17-24.
- [3] 吴春,王成荣,朱时继,等.BP神经网络在鸭西白垩系测井岩性识别中的应用[C]//2021油气田勘探与开发国际会议论文集(下册),2021:741-748.
- [4] 张昭杰,方石.基于遗传算法优化的支持向量机在岩性识别中的应用[J].世界地质,2019,38(02):486-491.
- [5] 黄体浩,李俊青,赵海勇.遗传算法优化的BP神经网络拷贝数变异检测[J].计算机工程与应用,2022,58(01):274-281.
- [6] 李曦,范翔宇,王兆峰,等.基于PSO-SVM的测井岩性识别方法研究——以南图尔盖盆地K油田古生界(P₂)储层为例[J].地球物理学进展,2022,37(02):617-626.
- [7] Pei-Yi Hao;Jung-Hsien Chiang;Yu-De Chen. Possibilistic classification by support vector networks[J]. Neural Networks,2022(149):40-56.
- [8] 于代国,孙建孟,王焕增,等.测井识别岩性新方法——支持向量机方法[J].大庆石油地质与开发,2005(05):93-95.
- [9] 同[6].
- [10] 杜陈宇,吴施楷.川中地区测井数据的支持向量机岩性识别[J].科技与创新,2018(08):124-125.
- [11] Friedl MA.,Brodley CE.DECISION TREE CLASSIFICATION OF LAND COVER FROM REMOTELY SENSED DATA[J].Remote Sensing of Environment:An Interdisciplinary Journal,1997,61(03):399-409.