

doi: 10.11720/wtyht.2018.0299
王嘉怡,郝立波,赵新运,等.基于土壤化学成分的基岩岩石类型识别[J].物探与化探,2018,42(6):1180–1185.<http://doi.org/10.11720/wtyht.2018.0299>
Wang J Y, Hao L B, Zhao X Y, et al. The identification of bedrock types based on soil chemical composition[J]. Geophysical and Geochemical Exploration, 2018, 42(6): 1180–1185. <http://doi.org/10.11720/wtyht.2018.0299>

基于土壤化学成分的基岩岩石类型识别

王嘉怡,郝立波,赵新运,马成有,陆继龙,赵玉岩,魏俏巧
(吉林大学 地球探测科学与技术学院,吉林 长春 130026)

摘 要: 浅覆盖区基岩露头少,地质填图精度低。根据岩石风化成土过程中的化学成分继承性,笔者提出了基于土壤化学成分识别基岩岩石类型的多层感知器神经网络模型。以大兴安岭北部阿龙山地区为例,根据火山岩基岩上覆土壤样品常量元素和亲石微量元素分析数据,有效地识别出了玄武岩类、安山岩类、英安岩类和流纹岩类 4 类基岩类型,识别的正确率达到了 90%。基于土壤化学成分识别基岩岩石类型的多层感知器神经网络模型具有方便、快捷、高效等优点,是提高浅覆盖区地质填图质量的有效途径之一。
关键词: 神经网络模型;地质填图;浅覆盖区;土壤;基岩
中图分类号: P632 **文献标识码:** A **文章编号:** 1000-8918(2018)06-1180-06

0 引言

我国浅覆盖区面积大,基岩露头少,在区域地质填图方面一直存在精度差、效率低等问题。基于水系沉积物对基岩化学成分的继承性,一些学者尝试利用水系沉积物的化学成分,采用聚类分析、因子分类等多元统计分析方法来判别基岩岩石类型^[1-6]。然而,这些分类方法多是基于某些相似性或差异性指标进行分类,选取不同的指标将出现不同的分类结果,因此分类结果的确定有很大的人为性。此外,来自不同基岩的水系沉积物在运移、就位等过程中往往会发生物质混合,也会干扰基岩岩石类型的识别^[7]。

地质系统的复杂性常导致地球化学信息表现出模糊性与非线性,而神经网络在处理该类问题方面具有明显的优势。相比于水系沉积物,残坡积土壤的化学成分受物质运移、混合等因素的影响相对较小。因此,采用残坡积土壤化学成分识别基岩类型也是提高浅覆盖区地质填图质量的有效途径之一^[8]。王大勇等利用 BP 神经网络模型对残坡积物

覆盖的未知地质体进行了识别,取得了良好的效果^[9]。郝立波等采用概率神经网络模型对 1:25 万多目标地球化学调查土壤数据进行了分析处理,在此基础上对第四系沉积物进行了识别分类,同样取得了良好的效果^[10]。为了研究残坡积土壤对浅覆盖区基岩类型识别的可行性,笔者系统采集了大兴安岭北部阿龙山地区不同类型的火山岩及对应的上覆残积土壤样品,尝试采用神经网络模型建立土壤化学成分与基岩类型的关系,识别基岩岩石类型。通过对多种神经网络分类模型效果的对比研究,提出了利用土壤化学成分识别基岩岩石类型的多层感知器神经网络分类模型,以期为提高浅覆盖区地质填图质量提供一种有效的方法。

1 方法原理

1.1 多层感知器神经网络模型

多层感知器神经网络模型由输入层、隐藏层和输出层构成(图 1),其中, $X_1, X_2 \cdots X_n$ 是输入; F 是非线性激活函数; $Y_1, Y_2 \cdots Y_n$ 是输出,可根据实际情况确定输出个数; $W_0, W_1, W_2 \cdots W_n$ 是权重;bias 是偏

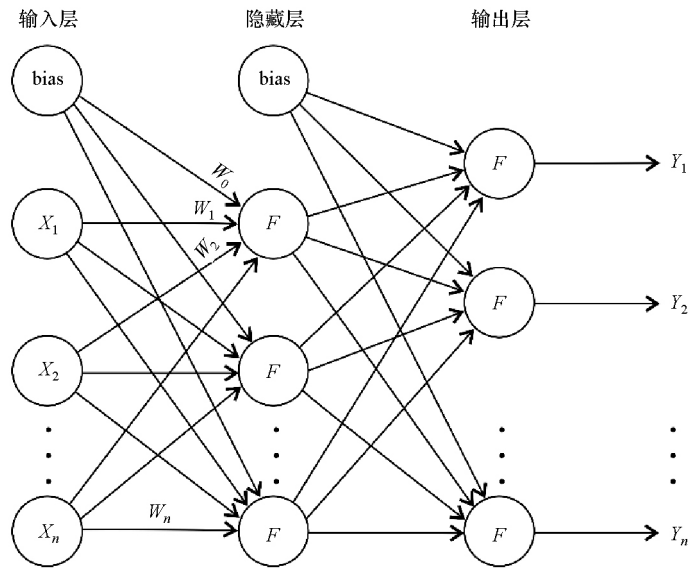


图 1 多层感知器神经网络结构示意图

置。权重在初始时是随机分配的,偏置通常设为 1。输入层、隐藏层、输出层之间连接的每一条线都有权重分配。

当样本和期望都已知时,称作有监督的训练模式,这时样本从输入层传入,经隐藏层传递到输出层。若实际输出与期望值不符,可将误差通过反向传播算法传递回输入层,调整权重。如此反复,直到符合期望,模型建立完成^[11]。本例中,隐藏层的激活函数为

$$F(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (-1 < F(x) < 1)。$$

式中 x 为输入值,经过公式处理后向下一层传递。激活函数种类繁多,如 Sigmoid 函数、双曲正切函数等,实际应用中可根据实际情况进行选择。地球化学数据间多存在非线性关系,激活函数可把非线性关系引入模型。

建立土壤化学成分与基岩岩石类型对应关系的多层感知器神经网络模型是将岩石类型已知的土壤样本作为训练样本进行学习,得到一个基岩岩石分类标准模型。标准模型的建立是基于神经网络的非线性分类能力。标准模型建立后,可将未知土壤样本数据输入模型进行准确识别,以确定所属岩石类型。

1.2 土壤样本数据归一化处理

为了消除土壤地球化学数据量纲不一致的影响,在进行神经网络运算之前,需要将数据进行归一化处理。数据归一化处理的公式为

$$y_i = \frac{x_i - x_{i,\min}}{x_{i,\max} - x_{i,\min}}。$$

式中: x_i 为土壤元素 i 的测量值, $x_{i,\min}$ 为所有土壤样

本中元素 i 的最小值, $x_{i,\max}$ 为所有土壤样本中元素 i 的最大值, y_i 为经过归一化处理后的值。

1.3 分类元素选取原则

分类元素应选择可以反映岩性变化(在不同岩石类型中含量差异大)、在表生作用中活动能力弱,且受矿化程度影响小的常量及微量元素。在实际应用中可选取常量元素和亲石微量元素,如 SiO_2 、 Al_2O_3 、 Fe_2O_3 、 FeO 、 CaO 、 MgO 、 Na_2O 、 K_2O 、 Li 、 Rb 、 Be 、 Sr 、 Ba 、 V 、 Ti 、 Sc 、 Cr 及 Zr 等。例如 Fe_2O_3 、 FeO 、 MgO 、 Cr 等富集于超基性岩中; CaO 、 Ti 、 V 、 Sc 等富集于基性岩中; SiO_2 、 Na_2O 、 K_2O 、 Li 、 Rb 、 Be 、 Sr 等富集于酸性岩中。

1.4 算法的实现

算法的实现主要在于建立土壤化学成分及其对应基岩岩石类型间的关系。首先,将土壤化学成分数据进行归一化处理;随后,选取一部分岩石类型已知的土壤样本作为训练样本,一部分作为预测样本;然后,将归一化后的训练土壤样本地球化学数据输入多层感知器神经网络模型中,通过神经网络的自我学习得到一个分类模型,当分类模型对训练样本的判别正确率较高时(一般在 95%),则认为建立的分类模型较为合理;最后,使用预测样本检验模型的泛化能力,当泛化能力较强时,可以判断模型是有效的。

2 实际应用

2.1 研究区概况

研究区位于大兴安岭以北的阿龙山地区(图 2),大地构造位置上属于中亚造山带东部额尔古纳

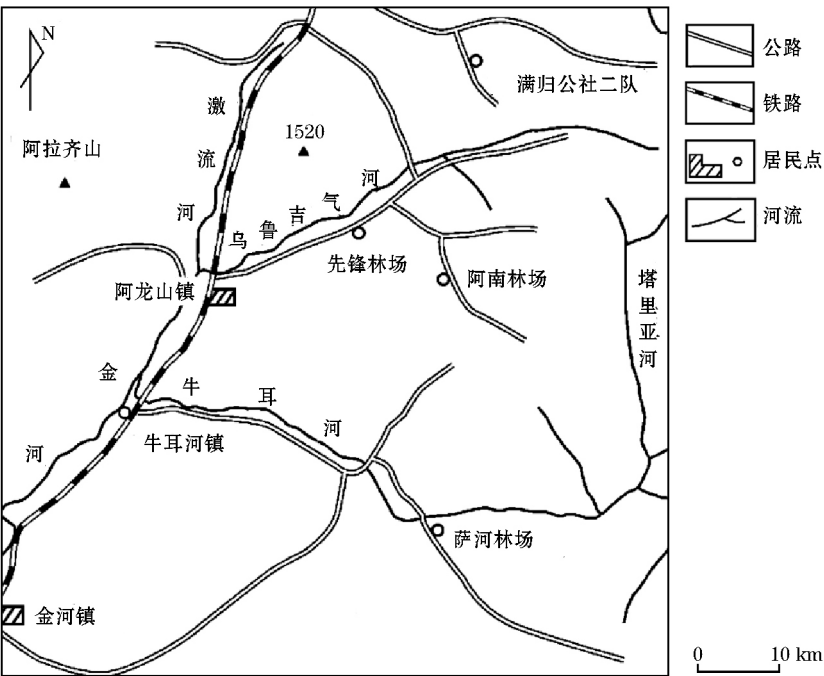


图2 阿龙山地区地理位置简图(据文献[12]修改)

地块与兴安地块的结合带,地势北高南低,属于中低山地形地貌。研究区气候属于寒温带大陆性气候,夏季短暂,温暖湿润,而冬季漫长,寒冷干燥。研究区内永久冻土较为发育。研究区内主要发育中生代火山岩,岩石风化以物理风化为主,化学风化较弱。火山岩风化形成的土壤中含有大量的石英、长石和云母等矿物。研究区覆盖严重,基岩露头少。

2.2 样品采集与分析

土壤样品及对应的基岩样品取自阿龙山地区公路附近的中生代火山岩分布区。样品采样点均布置于地形较平坦的山腰或半山腰处,以保证土壤样品为残积物。采用人工点槽方式采样,采集B层土壤样品167件及其对应下伏基岩样品167件。

土壤样品按照土壤地球化学测量规范处理,自然风干后,去除植物根系和岩石碎屑等杂物。岩石与土壤样品均粉碎至粒径小于200目(74 μm)。样品分析由吉林大学测试科学实验中心完成。分析项目共23项,包括常量元素SiO₂、Al₂O₃、TF₂O₃、K₂O、Na₂O、CaO、MgO和微量元素Ti、Mn、P、Cu、Pb、Zn、Co、Ni、V、Rb、Sr、Ba、Nb、Zr、Y、Th。氧化物和微量元素含量均采用粉末压片X射线荧光光谱法测试。主要分析步骤如下:用粉碎机将样品研磨至200目以下,经充分混匀后准确称取4.00 g,采用硼酸镶边垫底,在30 MPa压力下压制直径为40 mm的样片,然后采用X射线荧光光谱仪进行测试。测试过程中采用国家标准样进行质量监控。常量元素

分析相对误差小于5%,微量元素分析相对误差小于10%。

2.3 神经网络分类模型构建

2.3.1 分类元素的确定

在选取分类元素的过程中,首先剔除了可能受矿化影响的Cu、Pb、Zn等元素,然后,选择能够充分反映岩性变化的常量及微量元素。经多次实验,确定了SiO₂、TF₂O₃、CaO、MgO、Na₂O、K₂O、Ti、Co、Ni、V、Rb、Sr、Ba、Nb作为分类元素,并对其含量进行了归一化处理,作为神经网络分类模型的输入值。

2.3.2 激活函数的选择

在隐藏层中,常用的激活函数有双曲正切函数和Sigmoid函数。对比了这两种激活函数的应用效果,结果表明应用双曲正切函数时的样本预测正确率明显高于Sigmoid函数。因此,隐藏层中的激活函数选取了双曲正切函数。输出层的激活函数选择了适合于多分类问题的Softmax函数。

2.3.3 算法的选择

在分类模型的构建过程中,算法通常有标度共轭梯度法、梯度下降法和集成算法(如Boosting算法、Bagging算法等)。实验表明,集成算法在精度上明显优于标度共轭梯度法和梯度下降法,且以Boosting算法的精度最高,因此本文选择了Boosting算法。

2.3.4 构建模型及识别结果

由于岩石风化受气候、植被、水动力条件等因素制约,高寒气候条件导致大兴安岭地区物理风化作

用明显,化学风化作用较弱,土壤也没有发生明显的运移。因此,该地区岩石风化形成的土壤的化学成分对基岩有明显的继承性^[12-13],据此可以在土壤和基岩之间建立起非线性的对应关系。

首先,采用邱家骧硅—碱火山岩分类方法^[14]和岩矿鉴定,将基岩分为玄武岩类(玄武岩、玄武安山岩、碱性玄武岩等)、安山岩类(安山岩、石英安山岩等)、英安岩类(英安岩、英安流纹岩等)和流纹岩类(流纹岩、碱性流纹岩)4大类。基岩和对应土壤的分析数据统计结果分别列于表1和表2。

随机选取97件土壤样品作为训练样本,其余70件样品作为预测样本,其中,玄武岩类土壤训练样本15件,安山岩类土壤训练样本30件,英安岩类土壤训练样本22件,流纹岩类土壤训练样本30件。然后对挑选出的分类元素的含量进行归一化处理,并将归一化后的值作为多层感知器神经网络模型的输入值,采用Bootsing算法进行计算。计算结果表明,分类结果与实际岩石类型重合率为100%。

随后,将70件土壤预测样本数据输入建立的分类模型中,计算结果显示预测正确率达到了90%,其中玄武岩类正确率为100%,安山岩类为94%,英安岩类为100%,流纹岩类为84%,说明该神经网络

模型具有较强的泛化能力,可以用于浅覆盖区火山岩基岩岩石类型识别。

3 讨论

在多层感知器神经网络分类模型构建过程中,选择不同的分类元素,采用不同的激活函数和算法都会对模型的效果产生较大影响。笔者详细对比研究了不同分类元素、不同激活函数及不同算法对模型分类结果的影响。

在分类模型运算过程中,有些元素的缺失会影响分类效果,如CaO、TFe₂O₃等。去除CaO后,模拟输出重合率虽然也是100%,但模型对预测样本的判别正确率却只有63%,泛化能力较差,因此不宜舍弃CaO。对于TFe₂O₃,当不选取TFe₂O₃时,在分类中会把一个安山岩样本和一个流纹岩样本错误地识别为英安岩,使训练样本的模拟输出重合率降为90%,还会使预测样本的判别正确率降为60%,说明TFe₂O₃对样本的分类起重要影响,也不宜舍弃。而有些元素选用后反而会降低模型的分类能力,如Al₂O₃、Cu、Pb、Zn等。分析表明,Al₂O₃在各类岩石中的含量变化较小,在分类中没有起到重要的作用,

表1 阿龙山地区岩石分析结果统计

指标	玄武岩类(n=21)			安山岩类(n=47)			英安岩类(n=38)			流纹岩类(n=61)		
	最小值	最大值	平均值	最小值	最大值	平均值	最小值	最大值	平均值	最小值	最大值	平均值
SiO ₂	43.04	53.82	49.38	54.12	66.90	60.60	67.04	71.78	69.22	72.09	86.25	75.52
Al ₂ O ₃	10.21	19.24	15.95	13.82	18.78	16.63	12.64	17.83	15.83	8.77	16.03	13.26
TFe ₂ O ₃	6.36	20.50	10.26	3.23	9.54	5.81	1.85	6.83	3.23	0.96	3.31	1.79
K ₂ O	0.65	3.72	1.84	1.36	7.12	3.60	2.17	7.01	4.50	0.18	7.16	4.38
Na ₂ O	1.71	4.07	3.05	1.37	7.22	4.14	0.58	5.93	4.00	0.17	8.38	3.76
CaO	1.43	12.64	5.92	0.13	6.01	2.63	0.10	3.66	0.74	0.07	17.0	0.53
MgO	1.08	16.72	4.86	0.27	6.29	2.01	0.09	2.00	0.70	0.03	0.85	0.31
Ti	2248	29451	10552.7	2699	9704	5938.4	1460	4565	3170.2	474	3503	1536.3
Mn	555	3131	1404.3	402	2588	939.2	230	2052	809.5	93.8	1439	475.8
P	173	4922	2083.3	211	3155	1613.0	138	1514	726.7	62.4	797	242.9
Cu	0.52	63.6	22.2	0.48	69.6	14.8	0.70	8.63	21.6	0.18	13.1	2.9
Pb	2.85	31.0	11.7	1.69	80.8	20.4	9.08	50.3	21.6	5.71	201	28.6
Zn	59.7	100	103.6	34.2	143	82.5	42.0	159	67.3	18.4	105	49.1
Co	15.6	62.0	36.3	4.26	33.2	17.6	0.21	13.3	6.99	0.02	9.10	2.93
Ni	2.42	102	33.25	1.05	55.8	15.79	2.03	20.7	6.32	1.98	28.40	6.41
V	123	414	236.2	9.22	181	104.5	6.90	74.6	26.5	2.49	42.7	15.5
Rb	13.1	139	59.0	21.9	294	108.9	53.6	214	119.0	1.46	272	141.1
Sr	128	1525	570.1	20.2	1892	557.8	14.0	743	249.3	9.85	336	83.2
Ba	71.7	1793	625.0	122	2325	945.9	119	1655	933.6	48.5	1483	557.4
Nb	0.96	20.8	12.0	0.83	68.6	12.8	2.82	77.2	20.1	0.83	94.6	19.4
Zr	23.1	322	193.9	85.2	1649	286.2	190	1157	382.4	85.4	655	257.0
Y	5.81	43.5	20.6	7.87	74.4	17.8	9.1	99	25.5	3.34	129	19.4
Th	0.46	12.7	6.64	1.52	28.5	14.16	8.61	37.60	19.27	4.95	39.1	24.12

注:n为样本数;TFe₂O₃为全铁分析结果;氧化物质量分数单位为10⁻²,微量元素质量分数单位为10⁻⁶

表 2 阿龙山地区基岩上覆土壤分析结果统计

指标	玄武岩类(n=21)			安山岩类(n=47)			英安岩类(n=38)			流纹岩类(n=61)		
	最小值	最大值	平均值	最小值	最大值	平均值	最小值	最大值	平均值	最小值	最大值	平均值
SiO ₂	47.70	70.05	61.09	46.88	71.64	63.83	56.74	76.30	67.53	59.13	79.58	67.83
Al ₂ O ₃	13.89	18.82	15.90	12.08	18.50	15.52	10.10	17.37	14.77	11.03	18.85	15.15
TFe ₂ O ₃	4.45	11.33	7.37	3.80	8.69	5.65	2.18	10.38	4.86	2.32	6.71	4.27
K ₂ O	1.34	2.98	2.17	1.46	3.98	2.63	1.62	4.14	2.67	1.72	5.69	2.91
Na ₂ O	1.04	2.57	1.64	1.22	3.26	1.98	0.92	3.22	1.69	0.48	4.90	1.77
CaO	0.59	5.19	1.49	0.55	4.95	1.30	0.50	1.93	0.82	0.32	1.54	0.63
MgO	0.85	4.90	2.01	0.69	4.24	1.49	0.41	1.95	1.01	0.42	1.75	0.89
Ti	4811	11009	6908.5	4007	9939	5909.7	3803	7562	5457.9	1349	7901	4685.8
Mn	392	2515	966.0	269	2433	853.6	280	5625	919.9	219	4432	687.9
P	368	2971	864.7	315	3052	961.5	285	2560	768.7	152	1177	507.3
Cu	5.64	62.0	18.3	6.32	31.8	15.7	4.10	21.8	10.5	1.62	50.3	10.1
Pb	8.78	37.7	21.3	3.14	125	26.4	4.43	127	26.0	6.53	338	35.3
Zn	69.6	189	113.2	44.4	162	98.9	57.1	494	115.0	48.7	736	112.9
Co	10.4	34.3	21.9	8.8	60.5	18.2	6.33	22.5	13.0	0.54	17.7	10.1
Ni	2.56	62.4	21.5	2.51	33.9	17.2	4.60	1403	52.2	3.43	33.8	14.7
V	60.6	380	141.1	18.0	164.	97.6	31.5	128	75.2	14.4	113	65.1
Rb	62.8	143	103.8	52.1	201	116.3	11.0	165	118.1	61.2	221	128.8
Sr	119	431	233.0	141	884	306.5	66	384	178.5	34	268	131.6
Ba	471	704	600.7	475	964	658.3	444	1006	662.9	108	997	579.9
Nb	7.40	36.3	18.8	2.16	64.3	20.2	8.53	56.5	21.1	9.12	55.9	22.1
Zr	95.5	430	263.3	180	371	277.8	167	978	309.8	190	589	309.9
Y	5.52	45.0	21.2	9.85	41.2	18.9	12.8	54.2	22.6	12.6	43.2	21.1
Th	0.61	19.1	10.6	1.95	28.1	12.4	2.27	35.3	12.5	5.86	21.8	14.1

注:n 为样品数,TFe₂O₃ 为全铁分析结果,氧化物质量分数单位为 10⁻²,微量元素质量分数单位为 10⁻⁶

因此不宜选用。对于 Cu、Pb、Zn 等可能会受矿化影响的元素,会影响基岩分类,因此也不宜选用。

对于激活函数,虽然隐藏层中的双曲正切函数和 Sigmoid 函数都可以很好地解析数据中的非线性关系,但 Sigmoid 函数的输出不是 0 均值,收敛速度较慢。相比较而言,双曲正切函数的输出是以 0 为中心的,收敛速度较快。双曲正切函数的应用是对 Sigmoid 函数应用的一种改进。在实际应用中,双曲正切函数的效果多优于 Sigmoid 函数^[15-16]。笔者对比了两种函数的应用效果,结果表明两种函数的模拟输出重合率相差不大,但双曲正切函数的样本预测正确率明显高于 Sigmoid 函数 5%左右。

选择不同的算法,对模型的应用效果也有较大的影响。当选择 Boosting 算法时,模型的应用效果较好。该算法的基本思想是先用不同的分类方式对目标进行分类,形成多个基本模型,然后调整其中错误项的权重,最后将调整后的子模型整合形成分类模型^[17]。而其他算法的应用效果稍差,如标度共轭梯度法的模拟输出重合率为 67%,预测样本正确率为 71%;梯度下降法的模拟输出重合率为 81%,预测样本正确率为 76%。在集成算法中,Bagging 算法的模拟输出重合率也较高,达到了 91%,但其预测样本正确率却只有 64%。Bagging 算法为了保证采

样的均匀性,采取了随机取样方式,不同于 Boosting 算法的根据错误率取样方式。通常情况下,Bagging 算法虽然缩短了训练时间,但分类精度要低于 Boosting 算法^[18-19]。但在某些数据集中的样本中,采用 Bagging 算法的分类效果可能要好于 Boosting 算法,因此还需根据实际情况进行具体判断。

综上所述,在实际应用中应根据实际情况进行多次试验,确定最优分类元素、激活函数及分类算法,进而得到最优分类模型。

4 结论

- 1) 基于土壤化学成分建立的多层感知器神经网络模型对大兴安岭浅覆盖区火山岩基岩类型有很好的识别能力,利用土壤化学成分识别基岩类型是可行的。
- 2) 多层感知器神经网络模型具有训练快捷、使用方便、正确率高等优点,能较好地处理非线性对应关系,对提高浅覆盖区地质填图的精度有重要意义。
- 3) 不同激活函数、不同算法和分类元素的选择都会对多层感知器神经网络分类模型的效果产生影响。在使用过程中需要反复试验,确定最优分类模型。

参考文献:

[1] 刘德鹏,丁峰,汤正江. 区域化探在森林沼泽区地质填图应用初探[J]. 物探与化探,2004, 28(3): 209-212,217.

[2] 时艳香,纪宏金,郝立波,等. 利用水系沉积物地球化学数据判别浅覆盖区岩性与构造—欧氏距离法[J]. 物探化探计算技术, 2004, 26(3): 243-246.

[3] Ji H J, Zeng D M, Shi Y X, et al. Semi-hierarchical correspondence cluster analysis and regional geochemical pattern recognition [J]. Journal of Geochemical Exploration, 2007, 93(2): 109-119.

[4] 郝立波,陆继龙,李龙,等. 区域化探数据在浅覆盖区地质填图中的应用方法研究[J]. 中国地质,2007, 34(4): 710-715.

[5] 时艳香,郝立波,陆继龙,等. 因子分类法在黑龙江塔河地区地质填图中的应用[J]. 吉林大学学报:地球科学版,2008, 38(5): 899-903.

[6] 田密. 水系沉积物低弱地球化学异常提取方法研究[D]. 长春:吉林大学,2017.

[7] Zhao X Y, Hao L B, Lu J L, et al. Origin of skewed frequency distribution of regional geochemical data from stream sediments and a data processing method[J]. Journal of Geochemical Exploration, 2018, 194: 1-8.

[8] 徐国志,徐锦鹏,段玲玲. 化探资料在地质填图中的应用[J]. 物探与化探,2015, 39(3): 450-455.

[9] 王大勇,郝立波,陆继龙. 人工神经网络在识别浅覆盖区地质体中的应用[J]. 吉林大学学报:地球科学版,2006, 36(S2): 185-187.

[10] 郝立波,蒋艳明,陆继龙,等. 利用多目标地球化学数据识别第四纪沉积物类型——基于概率神经网络方法[J]. 吉林大学学报:地球科学版,2008, 36(6): 1081-1084.

[11] 蒋宗礼. 人工神经网络导论[M]. 北京:高等教育出版社,2001.

[12] 郝立波,陆继龙,马力. 浅覆盖区土壤化学成分与基岩化学成分的关系及其意义——以大兴安岭北部地区为例[J]. 中国地质,2005, 32(3): 477-482.

[13] 郝立波,马力,赵海滨. 岩石风化成土过程中岩石均一化作用及机理——以大兴安岭北部地区为例[J]. 地球化学,2004, 33(2): 131-138.

[14] 邱家骧. 岩浆岩岩石学[M]. 北京:地质出版社,1985.

[15] 李曦. 神经网络信息传输函数 Sigmoid 与 tanh 比较论证[J]. 武汉理工大学学报:交通科学与工程版,2004, 28(2): 312-314.

[16] Tavakolipour H, Mokhtarian M. Neural network approaches for prediction of pistachio drying kinetics[J]. International Journal of Food Engineering, 2012, 8(3): 1-15.

[17] 连克强. 基于 Boosting 的集成树算法研究与分析[D]. 北京:中国地质大学(北京),2018.

[18] Sharkey A J C. Boosting using neural networks[G]//Combining artificial neural Nets. London: Springer, 1999:51-78.

[19] 于玲,吴铁军. 集成学习:Boosting 算法综述[J]. 模式识别与人工智能,2004, 17(1): 52-59.

The identification of bedrock types based on soil chemical composition

WANG Jia-Yi, HAO Li-Bo, ZHAO Xin-Yun, MA Cheng-You, LU Ji-Long,
ZHAO Yu-Yan, WEI Qiao-Qiao

(College of Geo-Exploration Science and Technology, Jilin University, Changchun 130026, China)

Abstract: Geological mapping of areas with shallow overburden generally has less bedrock outcrop to work with, and is thus characterized by poor accuracy. Soils formed by weathering of rocks have significant inheritance of chemical composition from the bedrocks. In view of such a situation, the authors propose a soil chemical composition based multi-layer perceptron neural network model to recognize bedrock types. Taking Alongshan area in the northern part of the Da Hinggan Mountains as an example, the authors used geochemical data of major and lithophile trace elements of soil samples overlying volcanic bedrocks to identify bedrock types, and identified 4 types of bedrocks, i.e., basalt, andesite, dacite and rhyolite. The results show that the prediction accuracy of the model in the identification of bedrock types in the shallow overburden area of the Da Hinggan Mountains reaches up to 90%. The authors have reached the conclusion that soil chemical composition based multi-layer perceptron neural network model used for the identification of bedrock types has the advantage of high convenience, high speed and efficiency, and can provide an effective way for improving the geological mapping quality in areas with shallow overburden.

Key words: neural network model; geological mapping; shallow overburden area; soil; bedrock

(本文编辑:蒋实)