

# 矿井煤与瓦斯突出强度机器学习预测方法探讨

周松元<sup>1</sup> 罗渭<sup>2</sup> 黄炳<sup>1</sup>

(1. 湖南安全技术职业学院(长沙煤矿安全技术培训中心), 湖南 长沙 410151;

2. 杭州奇越数据科技有限公司, 浙江 杭州 310015)

**摘要** 本文通过对高瓦斯突出矿井利民煤矿“四位一体”防突资料的数据挖掘, 研究选用了机器学习方法中MARS算法, 构建了该矿井突出强度回归预测模型。只要取得并输入待预测地点对应的矿井最相关的突出预兆信息, 就可以随时预测计算出对应的突出强度, 达到防突措施程度“精准施策”目的, 以实现最佳的安全经济效果。

**关键词** 突出矿井 机器学习 强度预测模型 精准施策

**中图分类号** TD713 **文献标识码** A **doi**:10.3969/j.issn.1005-2801.2020.08.055

## Discussion on Prediction Method of Coal and Gas Outburst Intensity based on Machine Learning

Zhou Song-yuan<sup>1</sup> Luo Wei<sup>2</sup> Huang Bing<sup>1</sup>

(1. Hunan Vocational College of Safety Technology (Changsha Coal Mine Safety Technology Training Center), Hunan Changsha 410151;

2. Hangzhou Qiyue Data Technology Limited, Zhejiang Hangzhou 310015)

**Abstract:** In this paper, through the data mining of "four in one" outburst prevention data in Limin Coal Mine of high gas outburst mine, MARS algorithm in machine learning method is selected to construct the regression prediction model of outburst intensity in the mine. As long as the most relevant outburst precursor information of the mine corresponding to the predicted location is obtained and input, the corresponding outburst intensity can be predicted and calculated at any time, so as to achieve the goal of "precise strategy" of outburst prevention measures, so as to achieve the best safety and economic effect.

**Key words:** outburst mine machine learning strength prediction model precision strategy

煤与瓦斯突出强度直接决定灾变后果。经验表明: 100t 以下的低强度突出, 灾变区域可能仅涉及发生地点几十米局部范围内, 依靠通风设施, 就能迅速恢复正常, 造成的人员与财产损失轻微; 但数百吨、上千吨以上突出一旦发生, 其强烈冲击波会在极短时间推动灾变区域迅速扩大, 灾难性后果将是大概率事件。突出强度的预测具有非常重要的现实意义。

近来人工智能机器学习技术的迅猛发展, 给煤与瓦斯突出预测预警技术研究提供了新的视角、方

法和手段。此前一些相关的研究应用了诸如神经网络等算法进行分析, 取得了不错的效果, 但是这些研究采用的预测指标获取难度大, 样本量少, 且缺少独立验证检验, 构建的模型存在过度拟合的可能性, 实用性受限<sup>[1-2]</sup>。本文以湖南高瓦斯突出矿井利民煤矿为例, 采用机器学习方法, 探讨矿井煤与瓦斯突出强度预测模型构建的可行性与可靠性, 以期能为高危矿井防突管理精准施策, 提出一种灾害度评判的普适性参考方法。

## 1 突出强度研究机器学习方法

### 1.1 方法与流程

机器学习的核心思想是对历史数据进行分析获得知识, 并对新数据作出预测。探讨分析或将要服务的对象, 假定是已投产多年的煤与瓦斯突出矿井, 已发生过煤与瓦斯突出(或视同突出)事故, 长期的生产实践过程中, 已经积累了较多的与突出相关

收稿日期 2020-05-20

基金项目 湖南安全技术职业学院. 长沙煤矿安全技术培训中心《关于公布 2018 年度安全生产关键科学技术研究项目立项结果的通知》(湘安职院通[2018]96 号) 科研资助(项目编号 AK201802)

作者简介 周松元(1964-), 男, 焦作矿业学院采矿工程专业毕业, 矿山安全正高级工程师, 湖南省安全生产委员会专家, 中国煤炭工业安全科学技术学会瓦斯防治专委会委员, 主要从事矿山工程、隧道工程有害气体防治技术教学和研究。

的记录资料，如“四位一体”防突措施实施过程记录等。机器学习能自动对突出信息数据进行归纳、回归分析，从而获得规律，从理论上讲，能够利用

这些规律对未知数据（如突出趋势与强度等）进行预测，所以，本文就选用机器学习算法来构建矿井突出强度的预测模型。

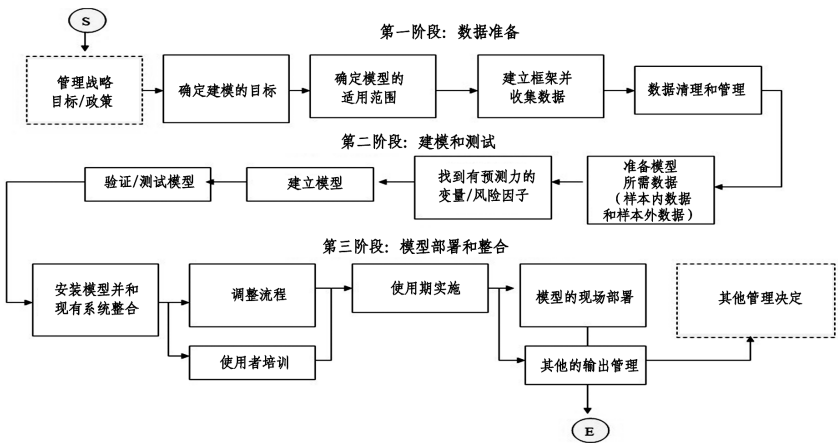


图 1 机器学习项目流程

采用机器学习方法进行完全的矿井煤与瓦斯突出强度预测研究，根据“CRISP-DM”跨行业机器学习标准流程，项目要经历三个阶段：数据准备阶段、建模与测试阶段和模型部署整合阶段（如图 1 所示）。对于探索性分析研究工作而言，本文探讨时段是到测试验证阶段。

1.2 算法选用

预测突出强度模型为回归预测模型。有许多机器学习算法适用于回归模型，评估回归模型的方法与经典统计学理论一致。在此主要使用  $R^2$  值进行模型评估。通过比较如 CART、TreeNet、随机森林和逻辑回归等各种算法建立的模型及运算结果比对，最终采用的是 MARS 算法建立的回归预测模型。

MARS（多变量适应回归样条）是一种高度自动化的回归分析工具，由斯坦福大学统计系 Jerome H. Friedman 开发<sup>[3]</sup>。MARS 可以自动选择使用哪些变量，对变量优化变形，发现变量之间的交互作用，并且模型能够自我测试，可有效防止诸如神经网络等算法可能出现的过度拟合问题。

MARS 的函数形式类似分段函数，但在经典统计中，分段函数切分位置通常被事先确定，并且被均匀间隔。在 MARS 中，切分位置则通过搜索步骤进行确认，并且只会保留需要数量的分段。MARS 创建了一组基函数来单独分解每个变量的信息，对于基函数运行回归等同于设定分段性线函数。

2 突出强度预测模型构建实例

2.1 数据采集分析

原国家统配煤矿总公司涟邵矿务局利民煤矿

（现政策关闭），根据矿井突出卡片记录及后期 10 年较为详细的“四位一体”防突指标测定记录及分析研究资料，进行数据挖掘整理得到突出样本 148 次，按建模经验以大约 1:8 的比例从正常“四位一体”综合防突措施记录中随机抽取 1180 条，合计 1328 条原始数据。由于时间跨度长，不同时期记录内容不完全相同，因此数据存在相当高的缺失率。统计情况如表 1 所示。

表 1 参与建模原始数据样本变量缺失情况统计

| 变量                            | 样本数量 | 缺失值数量 | 缺失值/% | 不同取值数量 |
|-------------------------------|------|-------|-------|--------|
| 样本序号                          | 1328 | 0     | 0.00  | 1328   |
| 突出强度（t）                       | 1328 | 0     | 0.00  | 67     |
| 波及距离（m）                       | 121  | 1207  | 90.89 | 44     |
| 瓦斯解吸指标 $K_1$                  | 1328 | 0     | 0.00  | 97     |
| 瓦斯解吸指标 $\Delta h_2$           | 647  | 681   | 51.28 | 19     |
| 最大钻屑量 S                       | 1328 | 0     | 0.00  | 13     |
| 瓦斯涌出初速度 Q                     | 36   | 1292  | 97.29 | 6      |
| 瓦斯涌出量（ $m^3$ ）<br>（突出前 8 小时内） | 96   | 1232  | 92.77 | 76     |
| 瓦斯浓度变化值（ $\Delta$ ）           | 1324 | 4     | 0.30  | 20     |
| 构造特征                          | 1328 | 0     | 0.00  | 8      |
| 煤层厚度变异系数                      | 132  | 1196  | 90.06 | 27     |
| 煤的坚固性系数值                      | 1317 | 11    | 0.83  | 17     |
| 最大开采深度（m）                     | 1322 | 6     | 0.45  | 85     |
| 温度变化                          | 1328 | 0     | 0.00  | 2      |

其中以突出强度是多少作为目标变量，其余指标作为预测变量，用于预测这个结果。由于部分预测变量缺失比例过高，而这些缺失通常是与工作流程高度相关的，即比如没有发生突出，“波及距离”就不会有数据记录。为了避免分析结果被这些因素干扰，将不使用相关指标进行分析。

## 2.2 模型构建过程

以往煤与瓦斯突出预测预警研究，研究样本或是采自不同矿区，或是同矿井，样本量采集数量通常也只有二三十条记录<sup>[1-2]</sup>，缺少独立验证的过程，因此可能存在过拟合的情况，即在实验室表现很好，但用于矿井具体生产过程时则偏差很大。为克服这一缺点，采用 20% 随机独立样本验证的方式，得到模型的  $R_2$  值为 0.428，其他表现指标如表 2 所示。

表 2 机器学习模型拟合结果评估指标

| 名称            | 训练样本            | 检验样本             |
|---------------|-----------------|------------------|
| RMSE          | 48.76849        | 158.29615        |
| MSE           | 2,378.36602     | 25,057.67047     |
| GCV           | 2,506.08257     |                  |
| MAD           | 15.83303        | 27.16477         |
| MAPE          | 3.85694         | 0.88774          |
| 目标变量平方和 (SSY) | 5,143,140.88867 | 11,259,549.64202 |
| 预测方差和 (SSE)   | 2,456,852.09649 | 6,439,821.37235  |
| $R^2$         | 0.52231         | 0.42806          |
| 正规化 $R^2$     | 0.52231         | 0.82297          |
| GCV $R^2$     | 0.49763         |                  |
| 调节 MSE        | 2,350.73737     |                  |
| 调节 $R^2$      | 0.51716         |                  |

构建的利民矿井煤与瓦斯突出强度预测模型（回归方程）为：

$$Y = 490.136 + 0.107283 \times BF1 - 0.0588377 \times BF2 + 133.792 \times BF3 - 0.0747148 \times BF5 - 0.205089 \times BF7 - 0.072001 \times BF8 - 2.22403 \times BF9 - 1.44556 \times BF10 + 68.0238 \times BF12 + 43.7664 \times BF13 - 60.5995 \times BF15.$$

预测模型中基函数形式则如下：

$BF1 = \max(0, \text{瓦斯涌出量}(M3) (\text{突出前 8 小时内推算量}) - 8186.4)$ ；

$BF2 = \max(0, 8186.4 - \text{瓦斯涌出量}(M3), \text{突$

$\text{出前 8 小时内推算量})$ ；

$BF3 = \text{构造特征 } \sin(\text{“断层、褶曲”}, \text{“穿层、褶曲”})$ ；

$BF5 = \max(0, \text{瓦斯涌出量}(M3), \text{突出前 8 小时内推算量}, -3050)$ ；

$BF7 = \max(0, \text{最大瓦斯解吸指标 } H2 - 460)$ ；

$BF8 = \max(0, 460 - \text{最大瓦斯解吸指标 } H2)$ ；

$BF9 = \max(0, \text{瓦斯涌出初速度 } Q - 61.365)$ ；

$BF10 = \max(0, 61.365 - \text{瓦斯涌出初速度 } Q)$ ；

$BF12 = \max(0, 1.01083 - \text{煤层厚度变异系数})$ ；

$BF13 = \text{工作面或钻孔温度有无变化 } \sin(\text{“有”})$ ；

$BF15 = \text{构造特征 } \sin(\text{“断层、褶曲”}, \text{“褶曲、穿层”})$ 。

## 2.3 预测模型验证

通过上面这个公式，只要取得了对应的上述指标值，就可以随时计算预测出对应的突出强度，就可以进行相应的预防工作，达到防突措施程度“精准施策”目的，以实现最佳的安全经济效益。下表为部分独立验证数据，依据上述公式可预测得到一个预测值，与实际值比较后发现预测误差一般在 50t 以内。如表 3 所示。

## 3 分析与结论

(1) 本次研究探讨了人工智能机器学习在煤矿瓦斯突出预测预警应用上的可能性，通过研究可以发现，即使在少量样本和数据质量不高的情况下，人工智能机器学习也可以较为准确地预测煤矿瓦斯突出强度，这说明人工智能机器学习在矿井地质灾难预测预警方面具有广阔的应用前景。

(2) 基于具体突出矿井的煤与瓦斯突出强度预测模型构建，可随时获得采掘工作面可能面临的灾变程度，有利于准确设计“四位一体”防突措施投入程度，防止投入不足或溢出，提高企业安全经济效益。

(3) 客观因素对预测分析精度以及适用性会造成一定影响。若数据时间跨度过长，实际工作流程也有可能已发生较大改变，矿井样本数据偏少和部分指标缺失率过高，也可能造成分析结果或具偶然性。

(4) 未来计划采用以下途径开展进一步研究：与具突出矿井的企业合作，采集大量规范准确的样本数据继续进行分析，进一步提升模型预测精度，并通过实践检验，最终实现将机器学习模型应用于煤与瓦斯突出强度预测工作。

表3 矿井煤与瓦斯突出强度预测验证表

| 突出强度预测值 /t | 强度预测偏差值 /t | 突出强度实际值 /t | 构造特征  | 温度变化 | 最大瓦斯解吸指标 / ( ml/g·min <sup>1/2</sup> ) | 瓦斯涌出初速度 / ( L/min ) | 瓦斯涌出量 /m <sup>3</sup> | 煤层厚度变异系数 |
|------------|------------|------------|-------|------|--|---------------------|-----------------------|----------|
| 130.771    | -32.7713   | 98         | 断层    | 有    | 360                                    | 4.335               | 1689.7                | 0        |
| 62.0108    | 15.9892    | 78         | 断层    | 有    | 360                                    | 4.335               | 1689.7                | 1.01083  |
| 43.9232    | 21.0768    | 65         | 断层    | 无    | 360                                    | 4.335               | 1689.7                | 0.633333 |
| 28.404     | 41.596     | 70         | 褶曲    | 无    | 460                                    | 4.335               | 1740                  | 3        |
| 80.9961    | 30.0039    | 111        | 断层    | 有    | 460                                    | 4.335               | 1890                  | 3        |
| 90.57      | -20.57     | 70         | 断层、褶曲 | 无    | 360                                    | 4.335               | 1200                  | 0.6      |
| 34.6742    | 35.3258    | 70         | 断层、褶曲 | 无    | 360                                    | 4.335               | 250                   | 0.6      |
| 37.6161    | 42.3839    | 80         | 断层、褶曲 | 无    | 360                                    | 4.335               | 300                   | 0.6      |
| 84.679     | -3.67905   | 81         | 褶曲、穿层 | 有    | 460                                    | 4.335               | 3300                  | 1.4      |
| 19.4388    | 35.5612    | 55         | 穿层    | 无    | 360                                    | 4.335               | 1710                  | 1.01083  |
| 69.2109    | -14.2109   | 55         | 断层    | 有    | 460                                    | 4.335               | 1689.7                | 1.01083  |
| 171.248    | 28.7518    | 200        | 穿层、褶曲 | 有    | 40                                     | 4.335               | 8186.4                | 3        |
| 84.9382    | -14.9382   | 70         | 断层    | 有    | 460                                    | 4.335               | 1957                  | 1.01083  |
| 102.178    | -48.1777   | 54         | 褶曲    | 有    | 460                                    | 4.335               | 2250                  | 2.2      |
| 55.6926    | 14.3074    | 70         | 褶曲    | 无    | 40                                     | 4.335               | 2000                  | 0.39     |
| 79.8639    | -19.8639   | 60         | 褶曲    | 无    | 360                                    | 4.335               | 4210                  | 1.4      |
| 165.293    | -22.2932   | 143        | 穿层、褶曲 | 无    | 240                                    | 4.335               | 8186.4                | 0.666667 |
| 33.5011    | 21.4989    | 55         | 褶曲    | 无    | 160                                    | 4.335               | 1950                  | 0.8      |
| 32.0151    | 27.9849    | 60         | 褶曲    | 无    | 360                                    | 4.335               | 1680                  | 0.8      |
| 103.74     | 16.26      | 120        | 穿层    | 无    | 100                                    | 4.335               | 4187                  | 0.39     |
| 148.772    | -18.7715   | 130        | 褶曲    | 有    | 460                                    | 4.335               | 3080                  | 1.01083  |
| 52.4905    | 47.5095    | 100        | 褶曲    | 无    | 100                                    | 4.335               | 1560                  | 0.12     |
| 79.0861    | 0.913902   | 80         | 褶曲    | 无    | 460                                    | 4.335               | 8186.4                | 0.2      |
| 100.172    | -48.1722   | 52         | 褶曲    | 无    | 600                                    | 4.335               | 2730                  | 0.39     |
| 117.547    | -44.5475   | 73         | 断层、褶曲 | 无    | 180                                    | 4.335               | 2110                  | 0.8      |
| 48.4315    | 28.5685    | 77         | 褶曲    | 无    | 360                                    | 4.335               | 1485                  | 0.39     |
| 15.7268    | 44.2732    | 60         | 褶曲    | 无    | 500                                    | 4.335               | 8186.4                | 1.01083  |
| 91.9707    | 5.02929    | 97         | 褶曲    | 无    | 260                                    | 61.365              | 8186.4                | 1.4      |

(下转第176页)



检修作业量,有利于加强对其他生产系统检修维护,提高生产系统稳定性。

### 3.4 经济效益增加明显

按照《煤业分公司商品煤内部收购管理办法》中“混煤、洗混煤收购价格表”计算:4607kcal 电煤价格 226 元/t;按照顾桥矿煤泥出矿价格表计算:煤泥价格 108 元/t;2019 年顾桥厂末煤系统吨末煤材料费、电费成本为 4.92 元/t,压滤系统吨煤泥材料费、电费成本为 22.0 元/t,根据 2019 年全年生产数据计算:

(1) 浅槽分选下限降低到 10mm 后,洗原煤量增加,压滤煤泥产量减少,因此火车商品煤、地销等收入增加,压滤煤泥收入减少,估算增加效益 1040 万元。

(2) 浅槽分选下限降低到 10mm 后,末煤系统入洗量和压滤煤泥产量减少,因此,末煤系统和压滤系统的生产成本降低,估算末煤系统材料费、电费成本降低 245 万元,压滤系统材料费、电费成本降低为 96 万元。

经以上分析,合计产生经济效益约 1381 万元。

(上接第 168 页)

度台,实现局/矿两级分级调度管理。

## 5 结语

物联网、智能装备是国家大力倡导的发展方向,本次应用研究是一个新技术集成、融合、创新的项目。

(1) 融合调度平台实现了调度通讯、人员定位、应急广播等多系统的统一调度管理,并实现集团统筹管控、矿井综合调度管理、作业现场实时调度指挥多级管控模式,提升了矿井综合管控水平。

(2) 基于 WiFi 的物联网综合基站,通过优化设计,对复杂工矿环境适应性强,具有功能全、成本低、升级方便的特点,为各类监测、监控系统提供统一的数据传输通道,解决煤矿井下设备多、通

(上接第 172 页)

### 【参考文献】

- [1] 王超杰,蒋承林,杨丁丁,等.煤与瓦斯突出强度预测研究现状分析[J].煤矿安全,2015,46(12):154-157.

## 4 结论

顾桥厂浅槽重介质分选机的分选下限降至 10mm 完全可行。调整浅槽分选下限后,与使用香蕉型原煤分级筛相比,不仅末煤系统开车时间缩短,提高了洗原煤量产率,降低了煤泥产量,优化了生产组织方式,减少了生产用工,同时为企业带来了 1381 万元经济效益。

### 【参考书目】

- [1] 黄亚飞,齐正义,徐春江,等.浅槽刮板重介质分选机流场试验研究[J].选煤技术,2012(06):27-32.
- [2] 姚学斌,史冰森.浅槽重介分选机分选下限和重介悬浮液上升流速度关系的研究[J].选煤技术,2019(02):8-13.
- [3] 王蒙.浅槽重介分选机常见问题与解决方案[J].选煤技术,2018(04):74-77.
- [4] 李世林.重介浅槽洗选生产中的注意事项[J].煤质技术,2005(05):13-15.
- [5] 陈志强,徐春江,黄亚飞.浅槽重介分选机内部流场的数值模拟研究[J].选煤技术,2012(01):1-4.

信网络繁杂、多系统重复建设的问题,可在各种类型的矿井推广应用。

(3) 智能终端的应用,在不增加井下作业人员负担的情况下,可实时掌握井下作业人员的地理位置、生命特征,提升个人安全防护水平,通过实时音视频通讯功能,提高矿井生产调度及应急指挥能力,提升安全生产效率。

### 【参考书目】

- [1] 高则海.矿井巷道无线信道特性分析与建模仿真[D].安徽理工大学,2016.
- [2] 刘增宝,边红星,吴相阔.基于 WiFi 技术的井下无线通讯系统在工业以太网上的应用[J].煤炭技术,2011,30(09):32-33.
- [3] 宋海鹰.煤矿井下无线通信系统的应用研究[J].山西煤炭,2017,37(03):74-76.

- [2] 史广山,王春光,高志扬,等.基于灰色-神经网络的新安煤田煤与瓦斯突出强度预测[J].煤矿安全,2015,46(09):166-168.
- [3] Friedman J H. Multivariate Adaptive Regression Spline[J]. Annals of Statistics, 1991, 19(01):1-67.