

变电设备的人工智能状态评价研究

王立德, 李晓武, 刘卫东, 李佐胜, 张超, 罗伟原

国网湖南省电力公司检修公司, 湖南长沙 410005

The Research on Condition Assessment for Substation Equipment Based on Artificial Intelligence

WANG Li-de, LI Xiao-wu, Liu Wei-dong, LI Zuo-sheng, Zhang Chao, Luo Wei-yuan
Maintenance Company of State Grid Hunan Electric Power Company, Changsha.

摘要: 变电设备状态评价一直依靠人工完成, 但人工评价效率低, 评价人员对规程规范的熟悉程度、评价尺度的把握不尽相同, 评价结果不尽相同。本文深入研究了传统方式下人工状态评价的方法、使用的数据、评价的流程, 提出了基于人工智能的状态评价方法, 并通过实例表明了该方法的有效性。

关键词: 变电设备; 状态评价; 人工智能; 神经网络; 支持向量机

ABSTRACT: The status evaluation of substation equipment is always depended on manual completion, but the efficiency of manual evaluation is low, the degree of familiarity with the rules and the grasp of the evaluation scale are different, and the evaluation results are different. This paper deeply studied the method of artificial state evaluation, the data used and the process of evaluation in the traditional way, and puts forward the state evaluation method based on artificial intelligence, which shows the effectiveness of the method through examples. Results indicate that the proposed method is effectively.

KEY WORD: substation equipment; condition assessment; artificial intelligence; neural network; support vector machine

1 前言

随着状态检修工作的开展, 每年需要对大量的电网设备进行状态评价, 并根据设备状态量评估设备的健康情况, 并给出检修决策建议, 该工作一直由人工开展, 随着电网规模的不断增长, 设备数量也日益庞大, 每年电网设备的定期评价工作需要大量熟悉设备状况的技术人员花费大量的时间完成。由于参与评价人员对规程规范的熟悉程度、评价尺度的把握不尽相同, 评价结果也存在标准不一, 评价存在误差的情况。

近年来, 人工智能技术不断进步, 在中文文本分析处理方面取得了很大的进展, 对电力设备状态信息进行综合分析可作为确定设备维修策略的有力依据^[1], 文献[2]中构建了基于卷积神经网络

的缺陷文本分类模型, 可显著降低分类错误率。文献[3]提出了提出基于深度语义学习的变压器运维文本信息挖掘方法, 通过对变压器的非结

构化文本数据进行信息挖掘, 有利于今后结合结构化数据, 用于评估变压器运行状态。本文将人工智能技术应用于变电设备的状态评价, 可大大减轻人工评价的工作量, 减少评价过程中评价尺度不一致的问题。

2 状态评价

2.1 状态评价流程

以某变电站变电设备状态评价信息为例对变电设备状态评价流程进行详细说明, 表 1-1 及表 1-2 为断路器的状态评价信息, 表中包含四部分信息:

(1) 设备基本信息。包含设备所属供电公司、变电站、间隔单元、电压等级、设备名称、设备型号、出厂日期、投运日期、生产厂家、设备类型。

(2) 设备检修信息。包含上次检修日期、

上次检修类别。

(3) 状态评价信息。包含本次评价日期、本次评价结果、当前状态描述。

(4) 检修决策建议。包含建议检修类别、建议检修时间、处理建议。

状态评价需根据收集到的设备状态信息按照设备类型选择对应的设备评价导则，将设备状态信息在“当前状态描述”按照设备评价导则对应条款对该设备健康状况进行打分，并生成“本次评价结果”。状态评价人员按照《变电设备决策导则》，根据设备基本信息、检修信息、评价

信息等进行综合评估，生成检修决策建议。具体评价流程见图 1。

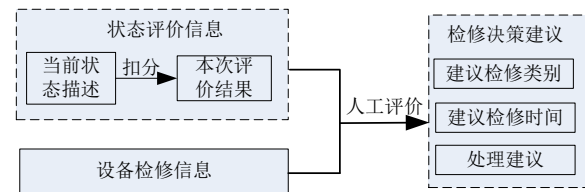


图 1 状态评价流程图

Fig 1 Flowchart of condition assessment

表 1-1 状态评价数据（一）

Tab 1-1 Data of condition assessment(1)

所属供电公司	变电站	间隔单元	电压等级	设备名称	设备型号	出厂日期	投运日期	生产厂家	设备类型
国网 XX 供电公司	220kVXX 变电站	#2 主变 620	交流 220kV	XX 变 620 断路器	LW10B-252W	2002/07/01	2002/12/01	XX 电气股份有限公司	断路器

表 1-2 状态评价数据（二）

Tab 1-2 Data of condition assessment(2)

上次检修日期	上次检修类别	本次评价日期	本次评价结果	当前状态描述	建议检修类别	建议检修时间	处理建议
2016/11/22	C 类检修	2019/05/10	异常状态	[XX 变 620 断路器] [操动机构][机构箱-密封]2002 年出厂产品，机构未进行过解体大修，密封件已超过使用寿命（厂家规定的使用寿命为 12 年）； [操动机构][机构箱-密封]机构箱结构设计（或制造工艺）存在缺陷，机构箱体易进水受潮。 [操动机构][其它]机构内接触器、继电器、辅助开关、限位开关、切换开关等二次元件，已运行 16 年，存在卡涩、接触不良、切换不到位等问题。	B 类检修	2021/11/22	上次 C 类检修时间 2016-11-22，建议 2020-11-22 开展 C 类检修，进行例行试验、检查、维护、清扫，结合停电更换断路器机构阀体密封组件，接触器、继电器、辅助开关、限位开关、切换开关等二次元件。评价人：XXX

2.2 状态评价文本内容

按照状态评价流程，变电设备的状态评价主要依据五个字段的数据：“上次检修日期”、“本次评价日期”、“上次检修类别”、“本次评价

结果”、“当前状态描述”。状态评价人员根据上述五个字段数据通过人工评价最终生成三个字段数据：“建议检修类别”、“建议检修时间”、“处理建议”。上述八个字段数据类型如表 2 所示。

表 2 状态评价数据类型

Tab 2 Data of condition assessment

字段名称	上次检修日期	本次评价日期	上次检修类别	本次评价结果
数据类型	日期型	日期型	文本型	文本型
字段名称	当前状态描述	建议检修类别	建议检修时间	处理建议
数据类型	文本型	文本型	日期型	文本型

按照状态检修导则分类标准,“建议检修类别”分为:“A类检修”、“B类检修”、“C类检修”、“D类检修”、“E类检修”五类,一般取前三类;评价结果按照设备的健康程度分为:“正常状态”、“注意状态”、“异常状态”、“严重状态”四类。

本文利用人工智能的方法,通过对历史状态评价数据进行训练,挖掘检修决策建议与状态评价信息及设备检修信息之间关系,实现变电设备的智能状态评价。

3 基于人工智能的检修决策建议模型构建

根据上文分析,检修决策模型有五个输入变量,三个输出变量(如图2)。三个输出变量的数据结构不同(见表2),与输入不完全存在对应的逻辑关系,输入与输出的逻辑关系如图3所示。

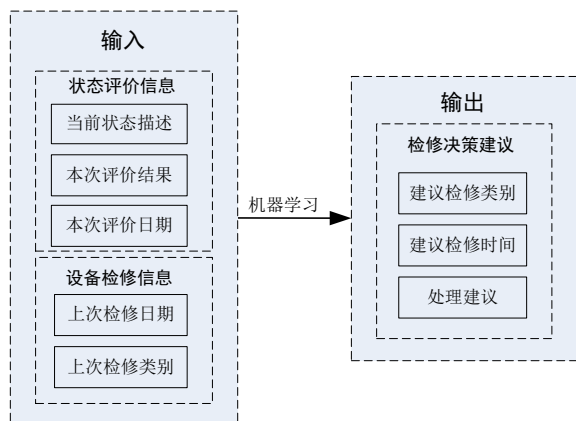


图2 检修决策建议模型

Fig 2 maintenance decision model

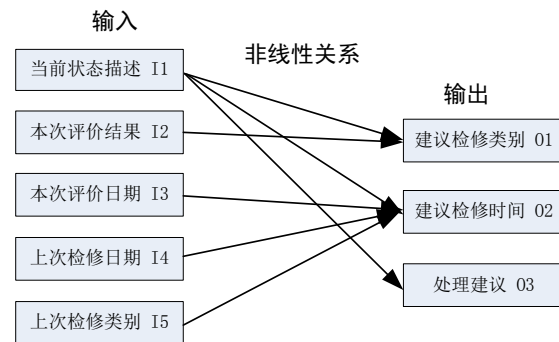


图3 输入与输出的逻辑关系

Fig 3 Logical relationship between input and output

3.1 输入/输出数据预处理

建立图3所示的基于人工智能的非线性关系模型需要将输入 $[I_1 \ I_2 \ I_3 \ I_4 \ I_5]$ 、输出 $[O_1 \ O_2 \ O_3]$ 数据进行预处理。

按照表2所示, I_1 、 I_2 、 I_5 、 O_1 、 O_3 为文本数据, I_3 、 I_4 、 O_2 为日期型数据。

3.1.1 文本型数据结构分析

以输出 O_3 为例,数据格式为“上次C类检修时间2016-11-22,建议2020-11-22开展C类检修,进行例行试验、检查、维护、清扫,结合停电更换断路器机构阀体密封组件,接触器、继电器、辅助开关、限位开关、切换开关等二次元件。评价人:XXX”。该文本可分为四个部分:

(1)第1部分:“上次C类检修时间2016-11-22”。为设备检修信息中“上次检修日期”和“上次检修类别”及连接词的组合。

(2)第2部分:“建议2020-11-22开展C类检修”。为输出变量“建议检修时间”和“检修检修类别”及连接词的组合。

(3)第3部分:“进行例行试验、检查、维护、清扫,结合停电更换断路器机构阀体密封组件,接触器、继电器、辅助开关、限位开关、切换开关等二次元件”。为纯文本描述语言,本文用向量 S 表示。

(4)第4部分:“评价人:XXX”。为评价人的基本信息。

通过上述分析,输出变量 O_3 的核心是第3部分文本 S 。其他文本 I_1 、 I_2 、 I_5 、 O_1 可以进行类似分解,并以向量形式进行表示。

3.1.2 日期型数据分析

以输出 O_2 为例,数据格式为“2021/11/22”,可以转换为数字格式 T ,用式(1)表示。

$$T = O_2 - I_4 \quad (1)$$

T 为建议检修时间 O_2 与上次检修日期 I_4 之间年份的差值,由于上次检修日期 I_4 已知, O_2 的求解可以转换年份差值 T 的求解。

3.1.3 输入/输出数据处理

建立基于人工智能的数学模型需将输入、输出数据进行预处理,并转变为计算机可以识别的形式,一般以矩阵或者向量的形式表示。

I_1 、 I_2 、 I_5 、 O_3 文本型数据的预处理,包括分段、分句、分词、去除停用词、标点符号等,并转变为向量。

O_1 虽然为文本型数据,但是考虑建模的方便(下文进行详细说明)以数字表示,“A类检修、B类检修、C类检修”分别用数字“1、2、3”表示。 O_2 经转换为 T ,也可用数字表示。

3.2 人工智能模型构建

图3所示的输入数据和输出数据之间为非线性逻辑关系,该逻辑关系可以用

$$[O_1 \ O_2 \ O_3] = f(I_1 \ I_2 \ I_3 \ I_4 \ I_5)$$

3.2.1 模型选择

(1) 输出 O_1 : 存在三种可能结果即:“A类检修”、“B类检修”、“C类检修”,可以用分类的方法处理,常用的分类器有朴素贝叶斯、贝叶斯网络、决策树、k-最近邻(kNN)、支持向量机(SVM)等人工智能方法来构建分类器,由于支持向量机(SVM)可解释性较好,是适合中小规模数据的最好的一类学习模型。本文选择采用支持向量机(SVM)模型。

(2) 输出 O_2 : 已转换为求解 T , 根据状态评价导则, T 取整数,通过对历史状态评价数据进行统计, T 一般取 0-12 之间整数,可转化为分类问题,同理采用支持向量机(SVM)模型。

(3) 输出 O_3 : 为文本型数据,对应的输入 I_1 同为文本型,可以采用神经网络模型,输入与输出之间的关系看作“黑盒”。

3.2.2 支持向量机(SVM)

支持向量机(SVM)是基于统计学习理论的人工智能方法。从本质上来说,它的基本的模型是特征空间上的间隔较大的线性分类器,学习的目的就是寻找类别间隔最大化的分类边界,最终可转化为一个凸二次规划问题的求解。SVM有可靠的理论依据,可解释型较强。

3.2.3 神经网络

神经网络是模拟人脑神经元相互作用进行信息处理的特性,由大量的神经元互相连接的网络。传统的神经网络模型如图4所示,根据处理问题的需要,逐步演化出了不同类型的神经网络结构,对于自然语言处理任务,序列到序列(seq2seq)模型是目前应用非常广泛的神经网络模型,可用于生成对话、机器翻译、文本摘要等,seq2seq模型结构如图5所示。

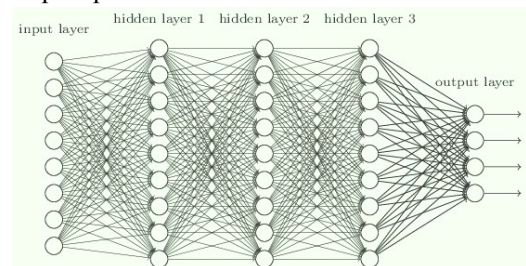


图4 神经网络模型

Fig 4 Neural network model

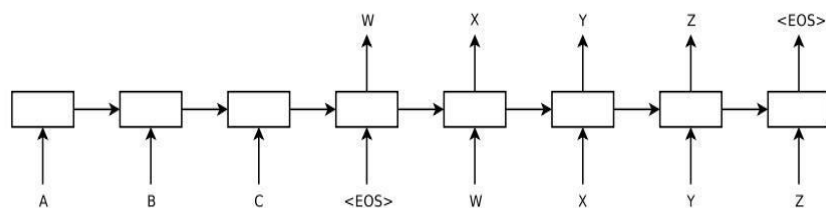


图5 seq2seq模型

Fig 5 seq2seq model

4 基于人工智能的状态评价实现

根据上文分析,检修决策模型有五个输入变量,三个输出变量(如图2)。三个输出变量的数据结构不同(见表2),与输入不完全存在对应的逻辑关系,输入与输出的逻辑关系如图3所示。

为研究本文所述方法对变电设备进行状态评价的效果,选取了2015-2019年某供电公司变电设备状态评价数据,由于部分设备状况良好,“当前状态描述”字段无不良情况的描述,该字段为空,故剔除该部分数据,其他数据作为样本数据集,共计1773条,分为训练集、验证集、测试集3份数据,其中70%数据为训练集数据,20%为验证集数据,10%为测试集数据,按照7:2:1的比例进行划分。

4.1 超参数设置

以计算“建议检修类别”为例,对SVM模型的超参数设置进行分析。SVM模型有3个主要超参数:

(1)惩罚参数 C : C 增大则对错误分类的惩罚增大,训练集测试时准确率提高,但泛化能力减弱。 C 减小,对错误分类的惩罚减小,容错能力提高,泛化能力较强。

(2)核函数 $kernel$: 可选的核函数包括“linear”、“poly”、“rbf”、“sigmoid”、“precomputed”。

(3)核函数系数 $gamma$: 该参数是 rbf、poly 和 sigmoid 的内核系数, $gamma$ 越大,则模型只能作用于支持向量附近,可能导致过拟合;反之, $gamma$ 越小,在训练集上分类效果不佳,可能导致欠拟合。

4.2 超参数调优

只有确定3个主要的超参数,方可对SVM模型进行优化计算。针对不同的问题,需要选择不同的超参数,由于本文所述SVM模型需要调整的超参数较多,需进行超参数调优,选择一组最佳超参数,常见的超参数调优方法主要有:

(1)网格搜索:在所有候选的参数中,通过循环遍历,寻找表现最好的一组参数,适用于三、四个(或者更少)的超参数。

(2)随机网格搜索:在一定的区间内,不断产生随机点,并计算其约束函数和目标函数的值,对满足约束条件的点,逐个比较其目标函数的值,将坏的点抛弃,保留好的点,最后便得到最优解的近似解。这种方法精度较差,找到近似最优解的效率高于网格搜索。一般用于粗选或普查。

SVM模型的主要超参数只有3个,故选择网格搜索法,其中参数选择范围设定为:

(1)惩罚参数 C : 1, 3, 5, 7, 9, 11, 13, 15, 17, 19;

(2)核函数 $kernel$: rbf, linear;

(3)核函数系数 $gamma$: 当核函数选择“rbf”时,需要设定该值,选择范围设定为 0.00001, 0.0001, 0.001, 0.1, 1, 10, 100, 1000;

通过计算,最终确定超参数最优取值为 $C=1$, $kernel='linear'$ 。此时训练集的计算精确度为100%,验证集的精确度为90.4%。

5 结论

随着电网规模的不断增长,每年大量的电网设备需要进行状态评价,通过对历史状态评价数据的大数据分析,并借助人工智能手段,可快速高效的开展状态评价工作,本文通过实例初步表明了该方法的有效性,人工智能状态评价的精准度还有待进一步研究。

参考文献

- [1]李刚,张博,赵文清,刘云鹏,等.电力设备状态评估中的数据科学问题:挑战与展望[J].电力系统自动化,2018,42(21):10-20
 - [2]刘梓权,王慧芳,曹靖,邱剑.基于卷积神经网络的电力设备缺陷文本分类模型研究[J].电网技术,2018,42(2):644-650
 - [3]蒋逸雯,李黎,李智威,等.基于深度语义学习的电力变压器运维文本信息挖掘方法[J].中国电机工程学报,2019,39(14):4162-4171
- 收稿日期:
作者简介:
王立德(1981-),男,河南省南阳市,硕士,高级工程师,从事变电设备状态检修。