

油水井计量仪表诊断分析与平台系统设计

王雄¹, 刘凯¹, 刘伟¹, 梁军¹, 罗大成¹, 石国亮¹,
曹泽宇², 侯增方²

¹荆州市明德科技有限公司, 湖北 荆州

²山东省微远科技有限公司, 山东 东营

收稿日期: 2023年3月16日; 录用日期: 2023年4月19日; 发布日期: 2023年4月26日

摘要

由于目前国内大数据分析技术正在起步阶段, 特别是在数据分析筛选失准(或故障)油水井计量仪表领域还没有实质性研究成果。为此, 开展了数据预处理模型、数据挖掘和失准(或故障)定位模型的研究, 并根据研究成果完成油水井计量仪表诊断分析平台的建立。通过研究, 建立了基于深度自编码学习的数据预处理模型; 建立基于神经网络数据分析的油水井计量仪表失准分析模型; 开发了基于神经网络分析的油水井计量仪表诊断分析平台。通过数据分析平台对胜利油田油水井计量数据的分析, 找出失准或故障传感器, 大大减少了传感器校准的工作量, 表明应用基于神经网络分析的油水井计量仪表失准分析模型是可行的, 具有科学性; 据此开发的油水井计量仪表诊断分析平台具有可用性。

关键词

油水井, 计量仪表, 诊断分析, 数据预处理, 数据挖掘, 数据分析平台

Oil Well Measuring Instrument Diagnostic Analysis Technology and Platform System Design

Xiong Wang¹, Kai Liu¹, Wei Liu¹, Jun Liang¹, Dacheng Luo¹, Guoliang Shi¹,
Zeyu Cao², Zengfang Hou²

¹Jingzhou Mingde Technology Co., LTD., Jingzhou Hubei

²Shandong Weiyuan Technology Co., LTD., Dongying Shandong

Received: Mar. 16th, 2023; accepted: Apr. 19th, 2023; published: Apr. 26th, 2023

文章引用: 王雄, 刘凯, 刘伟, 梁军, 罗大成, 石国亮, 曹泽宇, 侯增方. 油水井计量仪表诊断分析与平台系统设计[J]. 机械工程与技术, 2023, 12(2): 132-141. DOI: 10.12677/met.2023.122016

Abstract

At present, the domestic big data analysis technology is in its infancy; especially in the field of misalignment (or failure) of data analysis method screening oil and water Wells, there are no substantive research results. To this end, the research of data preprocessing model, data mining and misalignment (or fault) location model is carried out, according to the research results, to complete the establishment of oil and water well metering instrument diagnosis and analysis platform. Through the research, a data preprocessing model based on deep self-coding learning is established. Based on neural network data analysis, the misalignment analysis model of oil and well measuring instrument is established. A diagnostic and analysis platform for oil and well measuring instruments based on neural network analysis is developed. By analyzing the measurement data of oil and water Wells in Shengli Oilfield on the data analysis platform, the misalignment or fault sensors are found out, which greatly reduces the workload of sensor calibration. It shows that the application of the misalignment analysis model of oil and water Wells measuring instruments based on neural network analysis is feasible and scientific. Based on this, the diagnostic analysis platform of oil and well measuring instrument is available.

Keywords

Oil Well, Measuring Instrument, Diagnostic Analysis, Data Preprocessing, Data Mining, Data Analysis Platform

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着信息技术、计算机技术以及网络通信技术的发展,大量的数据采集及监控系统(SCADA)被运用在油气田的生产自动化管理中[1] [2]。如苏丹 1/2/4 区油田及其原油外输管道 SCADA 系统[3]。

信息技术的广泛应用离不开计量器具的作用,大量的计量器具的周期检定/校准成了信息系统准确运行的前提。近年来,由于信息化带来的计量器具数量的剧增,大部分监控类计量器具的在线校准开始采用数据分析技术快速定位疑似失准计量器具,然后再采用在线校准技术进行测试的方法。如英国石油公司(BP)应用节点数据信息故障检测、传感器故障交叉验证等算法[4],在巴西油区某采油厂,通过全网式的数据采集对无线感应器进行故障诊断。壳牌与惠普应用数据挖掘技术[5],在墨西哥油区进行计量数据准确性试验。

国内油气田自动化技术发展较为不均衡。大庆油田正在进行信息化油田推广,并且以单个生产过程自动监控系统居多,未实现整个油气田自动化管理[6]。长庆油田虽然信息化较早,但是标准不统一,并且低产和边缘油气井未实现信息化[7]。胜利油田“四化”推进快、普及广、标准化程度高,在国内处于领先水平[8]。但是,目前国内油田对信息化计量器具的检定/校准还比较传统[9]。同时,对计量仪表诊断技术的研究还比较少。汝继伟[10]针对传统油气水流量计量仪表检定装置存在系统能耗高、部分检验设备检验效率低、检验设备潜能未得到发挥、部分设备存在安全环保缺陷等问题,提出采取油气水流量计量仪表检定装置变频节能技术改造。李贵勇[11]介绍了油田载荷和位移计量仪表管理现状,给出了仪器校准/测试依据的标准和方式,对标准的适用性和校准方式量值传递的准确性进行了分析,并提出了相应的建

议。解咏炎[12]介绍了油田常用的高压注水计量仪表,对计量仪表设备的常见故障情况及其解决方法进行了分析。张思桐[13]认为油田计量仪表的智能化管理模式减少了数据处理中人工输入,提高了准确度,有利于后台数据维护、分析。房百庆[14]认为油田计量仪表在具体流量测量过程中,为了确保测量结果的真实准确性,可采取现场标定技术,为测量工作提供支撑保障。

通过调研,国内油田主要采用实验室送检和现场测试这2种方式。实验室送检主要采用检定/校准技术,优点是标准器准确度高、覆盖量程广;缺点是检定/校准数量有限,并且周期长、需备用表,同时安装、拆卸和调试费时费力。现场测试主要采用在线校准技术,优点是简便、快速;缺点是测试数量有限、只能覆盖部分量程,并且无法进行全量程调参,同时部分计量器具安装、拆卸费时费力。造成这一现象的原因主要是,信息化发展的程度不一,计量器具数量不同,测试需求也不同。同时,也与油田管理的要求有一定关系。以胜利油田为例,截止2015年已建成的67个“四化”项目的计量器具总量为110,597台(件);据估算,2017年底107个投产项目总量达到20万台(件);2019年起,胜利油田“四化”项目开始大规模进入运维阶段,研究一种简便、快速、准确性高的一体化计量仪表诊断测试技术,保证生产信息化系统中计量器具和数据传输的准确性十分有必要,也很迫切。这不仅能够维护生产信息化系统的准确运行,确保数字化油田的稳定生产,还能为采油管理区数据综合利用提供基础技术保障,为“四化”等相关部门大数据分析提供技术支持,为油田生产决策提供技术支撑[15]。

通过初步试验,得出数据分析方法能够筛选出失准(或故障)的计量仪表的初步结论[16]。但是,由于目前国内大数据分析技术正在起步阶段,通过文献检索和信息咨询,还没有企业、高校及科研院所在大数据分析筛选失准(或故障)油水井计量仪表领域有研究项目,更不用说实质研究成果。因此,应用数据挖掘技术研发筛选失准(或故障)油水井计量仪表的大数据分析软件势在必行。

为此,笔者以胜利油田为研究对象,开展了数据预处理模型、数据挖掘和失准(或故障)定位模型、数据降维和可视化模型的研究,并根据研究成果完成油水井计量仪表分析平台的建立。

2. 数据预处理模型

数据预处理(data preprocessing)是指在主要的处理以前对数据进行的一些处理。主要是清理异常值、纠正错误数据。现实世界中数据大体上都是不完整、不一致的脏数据,无法直接进行数据挖掘,或挖掘结果差强人意。为了提高数据挖掘的质量产生了数据预处理技术[17]。数据预处理有多种方法,包括数据清理、数据集成、数据变换、数据归约等。这些数据处理技术在数据挖掘之前使用,大大提高了数据挖掘模式的质量,降低实际挖掘所需要的时间。

对油井数据(包括回压、套压、温度、载荷、位移、电参数、电压、电流)和水井数据(包括压力、流量),在计量仪表类型、油水井工况记录、工况内容相关性等约束条件下,分别进行数据清理、数据选择、数据变换,处理不确定或不完整的数据和噪声,建立高支持度、高置信度的数据预处理模型。

2.1. 需要分析的数据内容

- 1) 每口油井半年以上的回压、套压、温度、载荷、位移、电功率、电压、电流分钟级数据。
- 2) 每口水井半年以上的压力、流量分钟级数据。
- 3) 相应时间段的油水井工况记录,包括开停井,注气注聚,仪表拆卸、更换等。
- 4) 关联数据的预测函数,包括油井载荷、位移、电参数、电压和电流的关联函数,水井流量和压力的关联函数。

2.2. 预处理模型

数据的预处理包括数据抽取、数据清洗、数据变换和数据加载。数据抽取就是从数据源中抽取跟本

主题相关的数据，本项目中数据是全抽取；数据清洗就是清除数据样本的部分离群噪声点、工况改变点及不完整数据点；数据变换能够将样本数据变换成统一的格式；数据装载就是将数据按统一格式进入分析系统中。该项目主要是进行数据清理。

数据清理例程通过填写缺失的值、光滑噪声数据、识别或删除离群点并解决不一致性来“清理”数据。主要是达到如下目标：格式标准化，异常数据清除，错误纠正，重复数据的清除。

2.2.1. 数据缺失

即某些数据的部分属性值缺失。如果对于某一属性值，缺失的数据占比很小，比如 10,000 条数据当中，缺失该属性的只有 10 条数据，那么我们可以简单去除这 10 条缺失数据。

也有一些缺失值可以用特定的内容来填充，例如：0，null，未知等，不过这个要根据实际来判断这样填充是否有利于后期的数据分析。

一些可推断的数值也可以用均值、中位数这些来填充，比如按照时间序列排序的数据。

2.2.2. 数据重复

重复这个问题也是需要数据做一个初步的判断，有些重复数据是无用的，需要查重并删除。但是有些情况，比如主题模型当中，需要统计词频，此时不但不能去重，而且相同词根的词也要做出处理，使其在统计时叠加；否则主题概率会出现问题。

2.2.3. 维度高

针对的是一些可以合并处理的属性或是对分类或聚类无影响的属性。

降维一般分为特征选择和特征检测 2 大方法。

2.3. 数据向量化

数据向量化是深度学习任务实现过程中最大的障碍之一。数据向量化是将数据转化为神经网络能够识别的格式。并且向量化的数据计算速度是非向量化数据的 250 倍，可以高效利用硬件算力资源，提高模型的运行速度。本项目向量化的数据有：

1) **油井功图数据**。因功图数据具有时序关系，所以将位移、载荷数据拼接为 1 条数据，形成以逗号分隔的字符串。将 3 个月的数据构建为 1 个 i 行 k 列 400 个时间序列的矩阵。

2) **油井温度压力等数据**。1 天 24 小时，每小时随机抽取 3 条数据共 72 条数据作为用于计算的 1 条数据，形成以逗号分隔的字符串。将 3 个月的数据构建为 1 个 i 行 k 列 648 个时间序列的矩阵。

3) **水井压力流量数据**。1 天 24 小时，每小时随机抽取 6 条数据共 144 条数据作为用于计算的 1 条数据，形成以逗号分隔的字符串。将 3 个月的数据构建为 1 个 i 行 k 列 288 个时间序列的矩阵。

得到以上矩阵后，经过数据缩放处理，处理为均值为 0，方差为 1 的值，防止单个大数据对于结果的影响过大。

3. 数据挖掘和失准(或故障)定位模型

对油水井数据，在工况限值、计量仪表性能，载荷、位移、电功率和电流耦合关系，压力和流量的耦合关系的基础上，分别进行特征化和区分、频繁模式、关联和相关性、分类和回归、聚类和离群点分析等多维数据挖掘，找出多维数据空间隐含计量仪表失准(或故障)的数据点。对上述数据点进行智能判断，定位故障计量仪表。定位模型的功能如下：

1) 主要依据载荷、位移构成的功图与电功率、位移构成的电功图的耦合关系，以及自相关性定位失准(或故障)载荷、位移。

- 2) 主要依据压力和流量的耦合关系, 以及自相关性定位失准(或故障)水井压力变送器和流量计。
- 3) 主要依据自相关和其他耦合关系, 定位失准(或故障)油井温度变送器和压力变送器。

3.1. 油井计量仪表分析模型

在基于 LSTM 网络和深度自编码算法中[18], 神经网络模型整体的输入参数为 400 个, 从而得出各层网络参数配置见表 1。其神经网络超参数及算法配置见表 2。

Table 1. Network parameters at each layer

表 1. 各层网络参数配置

网络层	输入参数	输出参数	网络类型
encoder0	400	100	LSTM
encoder1	100	80	LSTM
encoder2	80	5	LSTM
decoder1	5	80	LSTM
decoder2	80	100	LSTM
Output	100	400	RnnOutput

Table 2. Superparameters and algorithm configuration of neural network

表 2. 神经网络超参数及算法配置

名称	算法	值
权重初始化函数	XAVIER	
权重更新函数	RmsProp	学习率: 0.05
激活函数	TANH	
防止过拟合函数	L2	正则化系数: 0.0005
输出层损失函数	MSE	
输出层激活函数	IDENTITY	

对于油井井口实时数据, 神经网络模型整体的输入参数为 648 个(24*3*9, 即每小时取 2 次数据, 以 1 天的数据为 1 行), 从而得出各层网络参数配置见表 3。其神经网络超参数及算法配置同表 2。

Table 3. Network parameter configuration of each layer in real-time wellhead data

表 3. 油井井口实时数据各层网络参数配置

网络层	输入参数	输出参数	网络类型
encoder0	648	300	LSTM
encoder1	300	100	LSTM
encoder2	100	30	LSTM
decoder1	30	100	LSTM
decoder2	100	300	LSTM
Output	300	648	RnnOutput

3.2. 基于水井网络分析模型

对于水井井口实时数据,神经网络模型整体的输入参数为 288 个(24*6*2,即每小时取 6 条数据,以 1 天的数据为 1 行),从而得出各层网络参数配置如表 4 所示。其神经网络超参数及算法配置同表 2。

Table 4. Network parameter configuration of each layer in real-time wellhead data

表 4. 水井井口实时数据各层网络参数配置

网络层	输入参数	输出参数	网络类型
encoder0	288	300	LSTM
encoder1	300	100	LSTM
encoder2	100	30	LSTM
decoder1	30	100	LSTM
decoder2	100	300	LSTM
Output	300	288	RnnOutput

3.3. 完成各项模型可视化训练

图 1~3 为模型在可视化训练中的表现。可视化表现界面包含以下信息:

左上方——评分与迭代关系图,这是当前微批次的损失函数的值。

右上方——模型和定型信息。

左下方——所有网络权重的更新值与参数之比(各层)同迭代次数的关系。

右下方——激活函数、梯度和更新值的标准差(随时间变化情况)。

由图 1~3 可见,损失函数得分一直在平稳下降,并最终稳定在 0 分附近波动,证明此 3 个模型已经准确的提取了数据中的隐含特征,模型可以由此定型。

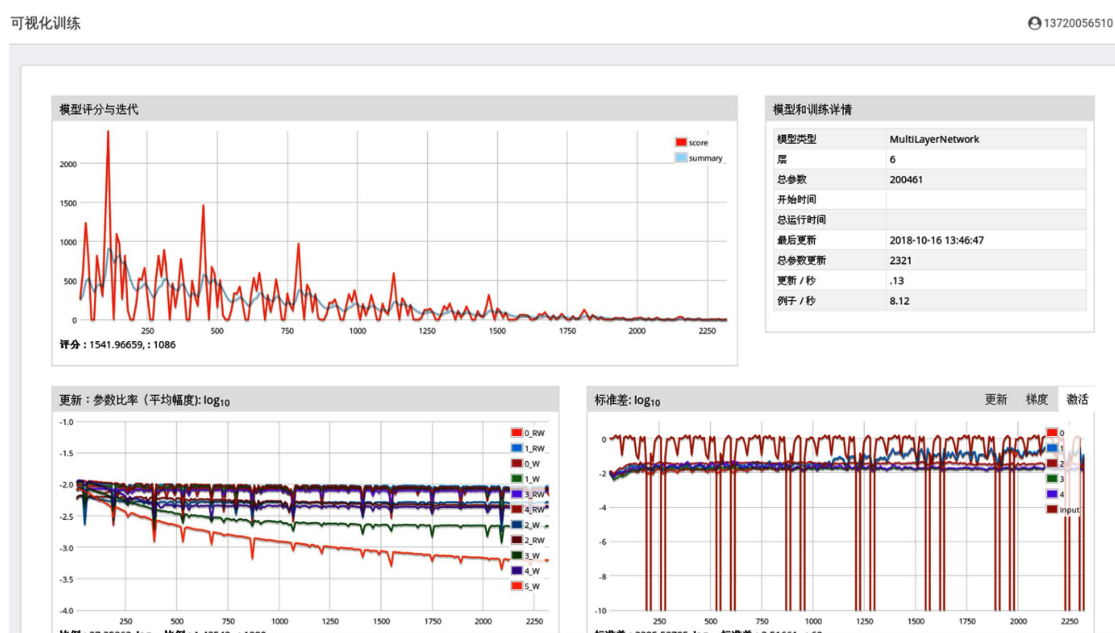


Figure 1. Visualization screenshot of oil well diagram model training

图 1. 油井功图模型训练可视化截图

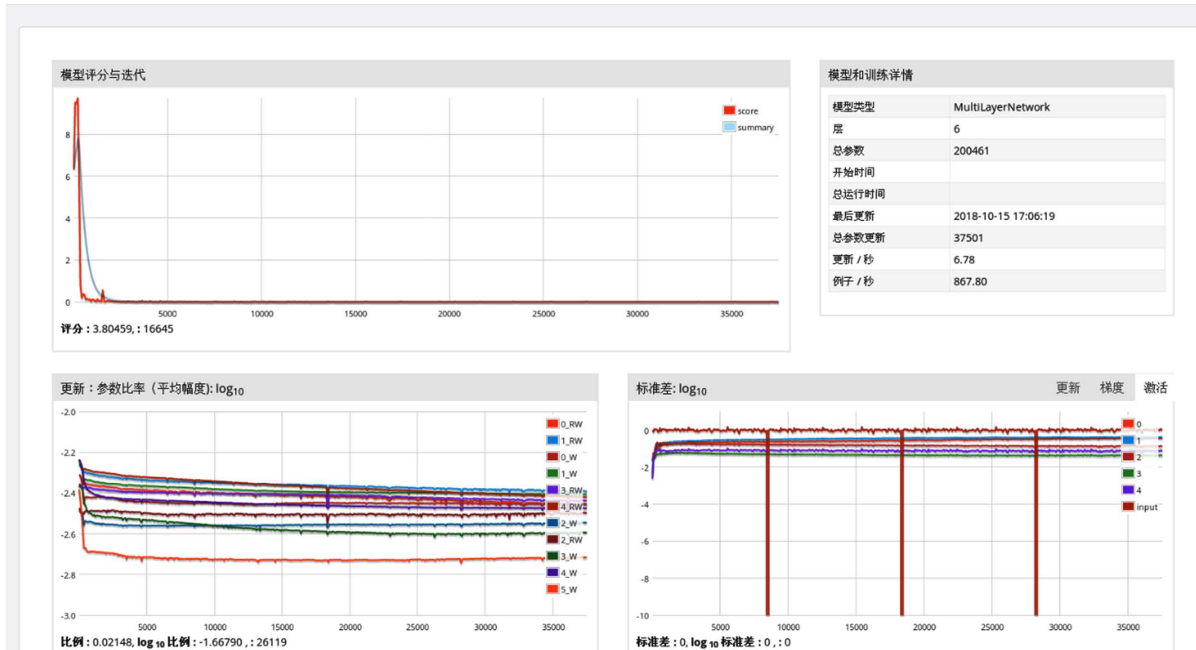


Figure 2. Visualization screenshot of oil well temperature and pressure model training

图 2. 油井温度压力模型训练可视化截图

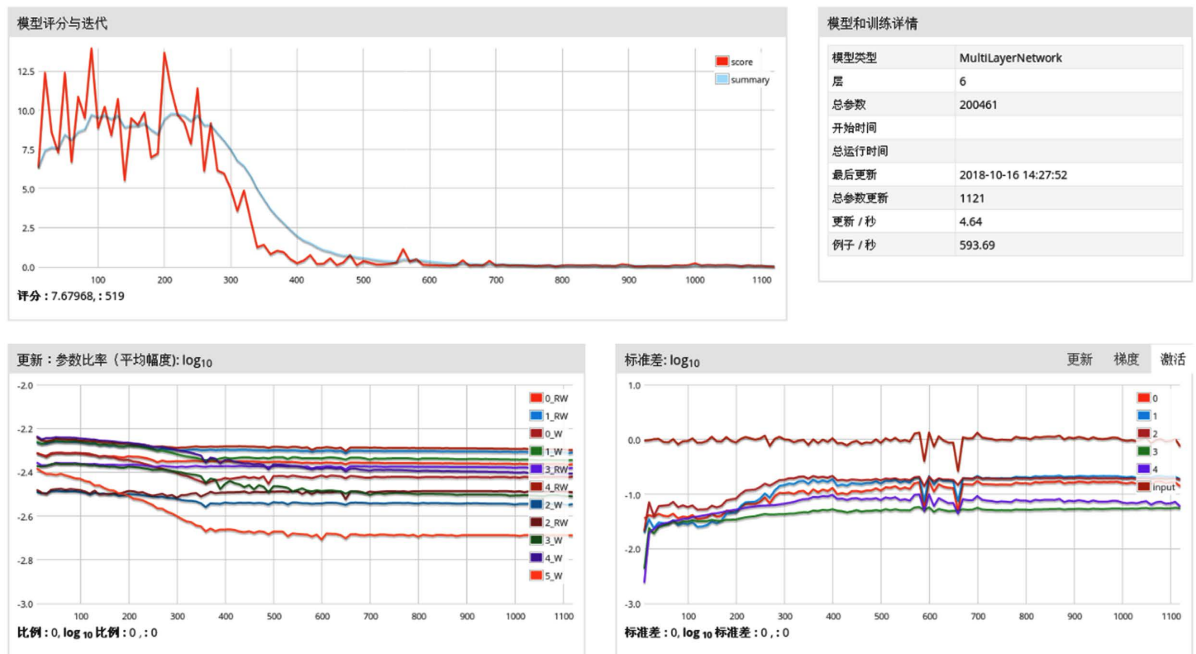


Figure 3. Visualization screenshot of well pressure flow model training

图 3. 水井压力流量模型训练可视化截图

在实际使用模型过程中，此模型还会在输出结果的同时，更新自己模型内的参数，达到自动进化的目的，可以让模型的表现越来越好。

4. 分析平台系统设计

油水井计量仪表诊断分析平台是一套综合性软件系统，部署在服务器上，前端通过浏览器进行数据交互和展示。

4.1. 主要功能

- 1) 数据输入。将信息化系统 SCADA 数据库数据导入到平台系统中。
- 2) 数据预处理。完成数据向量化和 ETL 处理，最后以标准化的数据进行学习。
- 3) 深度学习。平台系统的核心，集成机器学习的核心算法，经过大量数据训练，获取知识，建立模型。
- 4) 数据判别。对学习结果识别，进行传感器分类。
- 5) 数据展示。以图表方式展示结果。

4.2. 主要特点

- 1) 充分利用数据样本进行网络训练和模型拟合，无需故障样本数据，无需人工干预信号特征提取，即能实现对异常数据进行实时监测和判别。
- 2) 利用深度自编码器实现数据特征提取，最终通过损失函数得分来判定计量仪表的合格和待检测。
- 3) 全数据参与模型构建，在新数据加入后不断地学习，结果更准确。

4.3. 系统架构

基于人工智能能力平台，通过对数据质量模型、数据关联性模型及数据预测模型的研究，建立一种针对失准计量器具快速定位的数据模型，对现有油水井历史数据进行训练及分析诊断，以数据可视化展示，并提供数据报表查询及下载等功能[19]。系统构架框图如图 4 所示。

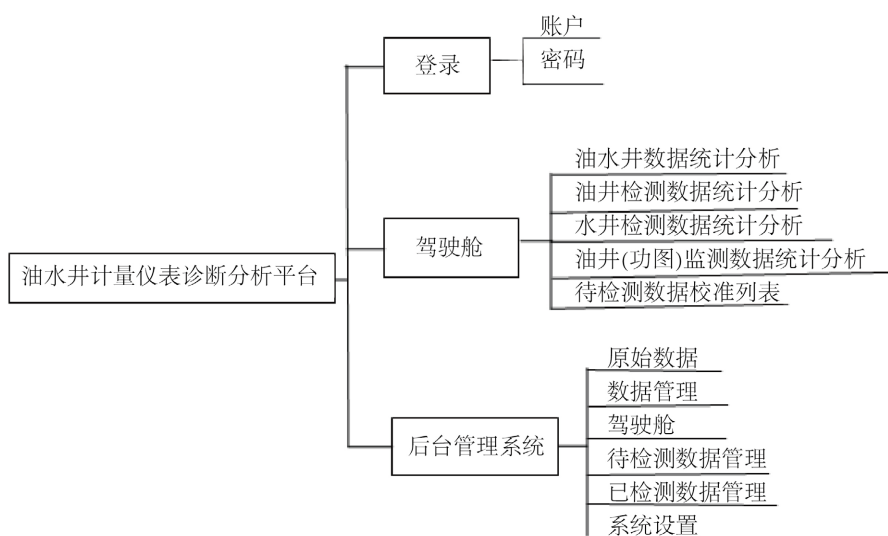


Figure 4. Block diagram of the system architecture

图 4. 系统构架框图

4.4. 系统运行

- 1) 系统初始化。油水井计量仪表诊断分析平台系统在初始化阶段对内置的管理员账号进行初始化，

包括对管理员登录 ID 初始化、登录密码初始化、个人信息初始化。

2) 运行控制。具体软件的运行模块组合为多个浏览器并发交互的运行环境, 各个模块在软件运行过程中能较好的交换信息, 处理数据。当用户登录到系统时, 用户输入的数据通过浏览器传输到服务器端, 由后台的管理模块对输入进行验证; 浏览器接收服务器返回的用户信息, 给不同的用户展示不同的界面。用户在浏览器上对学习信息进行操作, 提交数据给服务器后, 服务器校验数据; 服务器返回提交结果给浏览器, 是否修改成功。用户在浏览器使用油水井计量仪表诊断分析平台管理系统时, 能够见到漂亮清晰地界面, 简单的操作流程。

3) 运行结束。用户退出油水井计量仪表诊断分析平台管理系统后, 服务器会自动关闭与远程用户主机的连接。

5. 结语

1) 建立了基于深度自编码学习的降维数据预处理模型; 建立了基于神经网络数据分析的油水井计量仪表失准分析模型; 开发了基于神经网络分析的油水井计量仪表诊断分析平台。

2) 通过数据分析平台对胜利油田胜采、东辛和草东油水井计量数据的分析, 从而找出失准或故障传感器, 大大减少了传感器校准的工作量, 应用基于神经网络分析的油水井计量仪表失准分析模型是可行的, 具有科学性。据此开发的油水井计量仪表诊断分析平台具有可用性。

3) 因项目前期无法提供带有标签的数据, 所以目前设计的模型基于无监督学习的算法。本系统具有线下校准数据回填功能, 后期可以积累到一定数量的具有准确标签的数据, 可以根据准确的标签数据重新设计有监督学习算法模型, 对比有监督学习算法模型和无监督算法模型的预测结果, 选取最优模型作为最终交付物。

参考文献

- [1] 陈宏斌. 国际石油公司信息系统集成及启示[J]. 中国管理信息化, 2015, 18(23): 64-67.
- [2] 刘延明, 李建, 杨静. 基于 ESB 的石油行业系统集成的研究[J]. 信息技术, 2010, 34(2): 15-17.
- [3] 吴锐, 梁道君. 井口 SCADA 系统在苏丹 1/2/4 区的应用[J]. 科技资讯, 2008(26): 5-7.
- [4] 小晨编译. BP 如何利用大数据和人工智能改造油气行业[EB/OL]. 智慧能源之窗. <https://zhuanlan.zhihu.com/p/398293058>, 2021-08-10.
- [5] 大石油中的大数据: 壳牌使用分析技术推动业务成功的惊人方式[EB/OL]. <https://cda.pinggu.org/detail/3259.html>, 2018-11-15.
- [6] 常冠华, 陈树民, 熊华平, 等. 大庆油田油气勘探信息化建设的成果与特色[J]. 中国石油勘探, 2009, 14(5): 33-40+77.
- [7] 李建华. 长庆油田以信息化建设推进数字油田建设[J]. 数字石油和化工, 2006(12): 15-18.
- [8] 刘冲. 浅谈胜利油田四化建设与成果[J]. 石化技术, 2016, 23(9): 266+274.
- [9] 韩涛. 油田企业计量器具分类及检定校准研究[J]. 中国设备工程, 2019(24): 77-79.
- [10] 汝继伟. 油气水流量计量仪表检定装置节能措施探讨[J]. 石油石化节能, 2019, 9(1): 24-25+33+10.
- [11] 李贵勇. 油田载荷和位移计量仪表管理现状及建议[J]. 计量技术, 2019(8): 46-49.
- [12] 解咏炎. 油田高压注水计量仪表常见故障及解决办法[J]. 中国石油和化工标准与质量, 2021, 41(9): 20-21.
- [13] 张思桐. 浅析油田计量仪表管理存在问题及解决方案[J]. 中国设备工程, 2021(10): 144-145.
- [14] 房百庆. 流量计量仪表现场标定技术分析[J]. 中国设备工程, 2021(15): 149-150.
- [15] 胡斌. 胜利油田油公司改革模式探索[J]. 现代国企研究, 2019(8): 139.
- [16] 张芳. 计量器具在检定周期内失准的原因及应对办法[J]. 油气田地面工程, 2014, 33(2): 74.
- [17] 周泉锡. 常见数据预处理技术分析[J]. 通讯世界, 2019, 26(1): 17-18.

-
- [18] 刘鹏, 乔晓艳. 基于深度自编码和 LSTM 循环网络的脑电情感识别[J]. 测试技术学报, 2022, 36(2): 129-134.
- [19] 熊兆洪, 王春光, 李莉, 等. 一种油井计量仪器失准定位诊断方法、装置和终端[P]. 中国专利, CN111008203A. 2020-04-14.