

文章编号: 2095-2163(2020)05-0278-03

中图分类号: TP308

文献标志码: A

基于神经网络的数据中心故障预测方法研究

高林娥

(运城师范高等专科学校, 山西 运城 044000)

摘要: 作为社会基础设施的数据中心, 存储着大量的数据, 具有无法被替代的作用。文章以数据中心机房为研究对象, 从数据处理、模型设计、结果分析三方面, 围绕故障预测方法展开研究, 使神经网络的积极作用, 在故障预测过程中得到充分发挥。

关键词: 神经网络; 数据中心; 故障预测

Research on the Method of Data Center Fault Prediction Based on Neural Network

GAO Line

(Yuncheng Advanced Normal College, Yuncheng Shanxi 044000, China)

【Abstract】 As a data center of social infrastructure, it stores a large amount of data and plays an irreplaceable role. In this paper, the computer room of data center is taken as the research object, and the fault prediction method is studied from three aspects of data processing, model design and result analysis, so that the positive role of neural network can be given full play in the process of fault prediction.

【Key words】 Neural network; Data center; Fault prediction method

0 引言

科技的迅猛发展为智能化管理奠定了基础。研究表明, 将人工智能技术与深度学习相结合, 既能使数据中心所具有的智能化水平得到提升, 又能对数据中心运维效率产生的积极影响, 基于智能预警开展设计工作成为研究的重点。涉及机房故障诊断所开展的研究相关工作, 已经取得了良好的成果。但是, 基于神经网络的数据中心故障预测方法研究仍旧有待深入。

1 数据处理

数据中心机房(下文简称“机房”)往往拥有大量不同类型的设备, 导致故障出现的原因较多, 本文根据机房常见故障, 围绕故障预测的内容展开研究。利用神经网络良好的预测性能, 基于 TensorFlow 对神经网络模型进行搭建, 对经过归一化处理的数据日志数据的特点进行提取, 向训练模型进行传输, 使“基于设备运行数据, 对后续状态进行预测”的设想成为现实。

机房的基础设施主要有路由器、服务器和交换机等, 基础设施所表现出的性能, 往往会给网络运行效率带来直接影响。可以根据描述网络设备性能的参数指标, 包括 CPU 的利用率, 设备的响应时间, 报文的接收速率, 内存的利用率, 对设备的状态进行综合评价。但现有技术无法保证所采集日志数据,

不存在冗余数据和无效数据。因此有必要处理、加工原始数据。在完成数据采集工作后, 应该以设备状态为依据, 将样本数据分为正常和故障两类, 再将由此而获得的测试集数据、训练集数据, 用 csv 的格式进行保存^[1]。

本文以山西省某市自然资源厅的机房为研究对象, 通过测试、验证机房数据的方式, 获得最终结论。在此过程中, 所采用方法以主动轮询为主, 就是以所选取性能指标为依据, 对表示设备性能的数据进行采集。机房拥有大量不同类型的设备, 设备性能出现故障的原因, 也并不局限于某种或某几种, 本文只选取了具有代表性的故障类型, 例如: CPU 或内存的利用率过高, 致使设备有故障出现; 报文传输速率过低; 设备长时间未响应等等, 如表 1 所示。由于不同特征所对应量纲、量纲单位有所不同, 数据分析结果也会受到影响。为了消除量纲的影响, 有必要对数据进行标准化处理, 各项指标均处于相同的数量级, 再对指标进行对比和评价, 可以获得更加准确的结论。本文利用均方差法, 对数据进行归一化处理, 见公式(1):

$$Z = x_i - \mu / \delta. \quad (1)$$

其中, x_i 表示样本数据, 用来 μ 表示总体平均值, δ 表示总体标准差。经过归一化处理的数据, 满足正态分布曲线的特点, 即: 方差是 1, 均值是 0。这

作者简介: 高林娥(1976-), 女, 硕士, 讲师, 主要研究方向: 计算机应用技术。

收稿日期: 2020-03-13

也为后续对比、评价等工作的开展,奠定了基础。

表 1 样本数据
Tab. 1 Sample data

CPU/%	反应时间/m s	内用量/%	报文速率/ (数据包/s)
1.9	0.7	60.2	1.6
2.1	0.2	60.5	5.0
91.1	0.1	60.7	3.1
78.8	0.8	92.7	1.2

2 模型设计

2.1 模型概述

神经网络所指代内容,集中在生物领域和人工领域。其中,生物神经网络,主要是指由触点、细胞和大脑神经元所组成的,用来对生物意识进行产生的网络。人工神经网络,主要是指对动物神经网络所表现出的行为特征进行模仿,用来处理并行信息的数学模型,该网络主要通过对节点关系进行调整的方式,完成信息处理的任务。本文所讨论的神经网络,即为人工神经网络。在分类预测领域,神经网络表现出良好的性能。基于神经网络所设计的模型,主要由三层结构组成,输入层,隐藏层,输出层。需要选择相应的连接方式,对隐藏层进行连接,使模型深度得到扩充,提高模型所能解决问题的难度。数据中心机房内部设备众多,不同设备均有相应的性能,数据关系复杂,设备性能所呈现出的关系,以非线性关系为主,故可用于故障预测的神经网络,通常为多层神经网络。

2.2 研究方向

基于神经网络所展开研究涉及诸多领域,最具代表性的工作,体现在以下方面:(1)生物原型。从心理学、生理学等角度出发,对神经系统、网络核细胞所对应生物结构、功能机理进行研究。(2)建立模型。基于生物原型研究所取得成果,对理论模型进行建立,所建立的模型以数学模型、概念模型和知识模型为主。(3)算法。基于理论模型研究,对神经网络模型进行构建,用来模拟计算机或对硬件进行制作;与网络学习算法相关的内容,也可以纳入算法范围。

2.3 模型结构

选择利用 TensorFlow 框架,对神经网络模型进行搭建,为故障预测、训练等工作的开展提供帮助。TensorFlow 既可以作为算法学习接口,也可以作为基本框架,用来执行机器学习算法。在实际计算过程中,可将数据中心视为有向图,将计算操作视为节

点,将连接节点的位置视为边。以矩阵乘法、向量加法为代表的计算对故障进行预算的效果,往往十分准确。结构模型由输入层、隐藏层和输出层组成。输入层对数据进行输入,隐藏层对数据进行处理,输出层对分类结果进行输出^[2]。输入层所输入内容,以 CPU 的利用率,设备的响应时间,报文的接收速率,内存的利用率为主。隐藏层的操作,主要是矩阵相乘、激活函数,产生非线性特征。输出层所输出的状态,仅有正常和故障两种,用 0 表示正常,用 1 表示故障。另外,隐藏层的计算利用公式(2):

$$Z = f(x \cdot W + b). \quad (2)$$

其中, z 、 x 表示输出特征向量与输入特征向量, f 表示激活函数, W 表示权重矩阵, b 表示偏置。实验设置神经网络的层数为 2 层,第 1 层维度为 62,第 2 层维度为 124。

2.4 模型预测

将 SoftMax 函数视为对设备状态进行判断的主要概率计算,此函数的损失函数,多为交叉熵。在对模型进行训练的过程中,视情况对正则化项加以应用,避免出现过拟合的情况,提高预测的准确性。

2.5 评价指标

以预测机房故障为主,分类判断设备状态,属于典型的二分问题,可将分类的准确性、损失变化视为评测的主要指标,应用梯度下降法,使实验损失达到最小化。

3 结果分析

模型的计算流程如下:首先,构建计算图,例如,常量、初始化节点;其次,执行指点操作,检测常量、变量;最后,获得相应数据,完成故障的预测。将学习率设置成 0.01,利用梯度下降,将损失降到最低。训练次数为 500 轮。根据运行后准确率和损失率的具体数值,不断增加迭代次数,使模型损失曲线的趋势逐渐向收敛靠拢,最终趋于稳定^[3]。准确率和迭代次数的关系为正相关,也就是说,如果迭代次数增加,模型准确率也会有所增加,待准确率稳定后,将测试数据输入其中,能够对数据和设备状态的关系,进行相对准确的预测和呈现。

4 结束语

基于 TensorFlow 搭建模型,在经过系统训练后,具有以设备数据信息为依据,对设备状态进行预测的能力。如果设备呈现出的状态欠佳,可及时提醒工作人员,通过维护设备,减少发生故障的可能性。另外,实现智能预警、解决故障等方面,设备预测的价值也得以体现。(下转第 283 页)