

# 基于对抗生成网络与BP神经网络的低压台区 线损率预测

方舟<sup>1</sup>, 裘炜浩<sup>1</sup>, 季超<sup>1</sup>, 夏鹏飞<sup>2</sup>, 龚康家<sup>2</sup>, 周后盘<sup>2</sup>

(1. 国网浙江省电力有限公司信息通信分公司, 杭州 310007;  
2. 杭州电子科技大学, 杭州 310018)

**摘要:** 线损电量产生于发电、输电、配电及用电等环节, 线损电量占供电量的百分比称为线损率。针对线损率预测问题, 提出了一种基于对抗生成网络与BP神经网络的低压台区线损率预测模型。选取某市低压台区线损数据作为实验数据集, 经数据预处理后, 通过K-Means++算法将低压台区分类, 对不同类别的低压台区分别训练对抗生成网络来增加不同类别的样本数据, 利用不同类别样本数据分别训练BP神经网络搭建低压台区线损预测模型。实验结果表明, 与传统BP神经网络模型相比, 该预测模型具有更加准确的效果, 通过对抗网络增加样本数据可以有效改善低压台区线损数据量偏小的问题。

**关键词:** 线损率预测; 对抗生成网络; BP神经网络; K-Means++算法

文章编号: 1007-1881(2019)10-0046-06

DOI: 10.19585/j.zjdl.201910008

中图分类号: TM714

文献标志码: B

开放科学(资源服务)标识码(OSID):



## Prediction of Line Loss Rate of Low-voltage Transformer Area Based on Generative Adversarial Networks and BP Neural Network

FANG Zhou<sup>1</sup>, QIU Weihao<sup>1</sup>, JI Chao<sup>1</sup>, XIA Pengfei<sup>2</sup>, GONG Kangjia<sup>2</sup>, ZHOU Houpan<sup>2</sup>

(1. State Grid Zhejiang Information & Telecommunication Company Branch, Hangzhou 310007, China;  
2. Hangzhou Dianzi University, Hangzhou 310018, China)

**Abstract:** The line loss energy consumption comes from power generation, transmission, distribution and power consumption, and the percentage of the line loss power in the power supply is called the line loss rate. Aiming at the problem of line loss rate prediction, this paper proposes a low-voltage transformer area line loss prediction model based on the generative adversarial networks and BP neural network. The line loss data of a low-voltage transformer area in a city is selected as the experimental data sets. After data pre-processing, similar low-voltage transformer areas are classified by K-Means++ algorithm, and different types of low-voltage transformer areas are trained to increase different types of data by different generative adversarial networks. Finally, different types of sample data are used to train BP neural network and construct a low-voltage transformer area line loss prediction model. The experimental results show that the prediction model proposed in this paper is more accurate than the traditional BP neural network model, and it shows that increasing the sample data by generative adversarial networks can effectively improve the problem of small line loss data quantity in the low-voltage transformer area.

**Keywords:** line loss rate prediction; generative adversarial networks; BP neural networks; K-Means++ Algorithm

## 0 引言

线损率是电力企业一项重要的综合性技术经济指标, 不但能反映电网结构和运行方面的合理性, 而且可以反映电力企业的技术和经营管理水

平<sup>[1]</sup>。低压台区作为电力系统的末端环节, 快速准确预测低压台区线损率, 为制定合理的降损措施提供依据是供电企业的重要任务<sup>[2]</sup>。在实际电力生产中, 影响线损率的因素是复杂的, 这给线损率预测带来一定难度。因此, 如何建立一个合

理有效的线损率预测模型值得进一步研究。

为了建立一个合理有效的线损率预测模型,学者们进行了许多研究,提出了各种预测模型。文献[3]采用核偏最小二乘法进行线损率预测,核偏最小二乘回归可以看做是多元线性回归分析、典型相关分析、主成分分析三者的结合,该预测模型适合特征值较少的方法,特征值较多使该预测模型计算量过大。文献[4]采用GM(1,1)灰色模型,能够在历史数据不多的情况下,找到线损率预测的内在规律,但它只在影响电力系统线损的因素不出现很大变化的前提下适用。多元线性回归法在实际应用中,由于选择的变量之间往往存在很强的相互关系,导致其在建模精度和回归精度上都存在一定的不足。文献[5]提出了一种基于快速独立成分分析和支持向量回归的计算方法。智能算法的一个不可避免的问题是特征的选择,而特征多为有功功率、无功功率、线路长度、配电变压器容量等,没有综合考虑上月线损率、同期线损率、同期累计线损率等历史线损数据对线损的影响。

## 1 相关技术介绍

### 1.1 K-Means++算法

聚类是一种无监督的学习,即聚类是不需要类别的标注,通过聚类可以实现在大量数据中归类相似的数据<sup>[6]</sup>。最常见的聚类方法是K-Means算法,而K-Means++算法是基于K-Means算法改进的,K-Means++在选取第1个聚类中心时是随机产生的,与K-Means算法相同,但是在选取第 $n+1$ ( $n$ 是正整数)个聚类中心时,K-Means++算法是选择距离当前 $n$ 个聚类中心最远的点做为第 $n+1$ 个聚类中心,可以说这也符合我们的直觉:聚类中心当然是互相离得越远越好<sup>[7]</sup>。这个改进虽然直观简单,但是却非常有效。

K-Means++算法中的K表示的是聚类为K个簇,Means代表取每一个聚类中数据值的均值作为该簇的中心,或者称为质心,即用每一类的质心对该簇进行描述。通过计算各个K值对应的轮廓系数来确定最优聚类数 $K$ <sup>[8]</sup>。轮廓系数计算公式:

$$s(i) = \frac{b(i) - a(i)}{\max[a(i), b(i)]} \quad (1)$$

式中: $a(i)$ 是样本 $x$ 与簇内的其他点之间的平均距离; $b(i)$ 是样本 $x$ 与最近簇中所有点之间的平均距离。

由轮廓系数计算公式可分析得 $s(i)$ 接近1则说明样本聚类合理,则选择轮廓系数接近1的聚类数为最优聚类数 $K$ 。

### 1.2 对抗生成网络

GAN(对抗生成网络)是一种深度学习模型,是近年来最具前景的无监督学习之一<sup>[9]</sup>。对抗生成网络中包含两个网络:生成网络和判别网络,对抗生成网络结构如图1所示<sup>[10]</sup>。

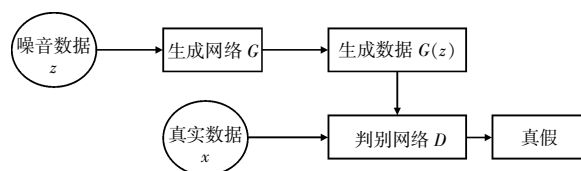


图1 对抗生成网络结构

生成网络 $G$ :不断学习训练集中真实数据的概率分布,目标是将输入的随机噪声数据转化为可以以假乱真的数据。判别网络 $D$ :判断一个数据是否是真实的数据,目标是将生成网络 $G$ 产生的“假”数据与训练集中的“真”数据分辨开<sup>[11]</sup>。训练过程中通过相互竞争让这两个网络同时得到增强, $G$ 和 $D$ 构成一个动态的“博弈过程”<sup>[12]</sup>。对于对抗生成网络来说,其判别网络、生成网络的损失函数分别定义为:

$$L_D = \log[D(x)] + \log\{1 - D[G(z)]\} \quad (2)$$

$$L_G = \log\{D[G(z)]\} \quad (3)$$

式中: $z$ 为噪音数据; $x$ 为真实数据; $D$ 为判别网络; $G$ 为生成网络。

随着时间的推移生成网络和判别网络不断进行对抗,最终两个网络达到一个动态平衡:生成网络生成的数据 $G(z)$ 接近于真实数据分布,而判别网络识别不出生成网络生成的数据 $G(z)$ 。最后,我们就可以使用生成网络来生成数据。

### 1.3 BP神经网络

BP神经网络是由Rumelhart和McClelland为首的科学家提出的概念,它是一种按照误差逆向传播算法训练的多层前馈神经网络<sup>[13]</sup>。BP神经网络的过程主要分为两个阶段,第一阶段是信号的前向传播,从输入层经过隐含层,最后到达输出层;

第二阶段是误差的反向传播,从输出层到隐含层,最后到输入层,依次调节隐含层到输出层的权重和偏置以及输入层到隐含层的权重和偏置<sup>[14]</sup>。

## 2 构建低压台区线损率预测模型

本文结合低压台区历史线损数据,提出了一种基于对抗生成网络与BP神经网络的低压台区线损率预测模型,如图2所示。

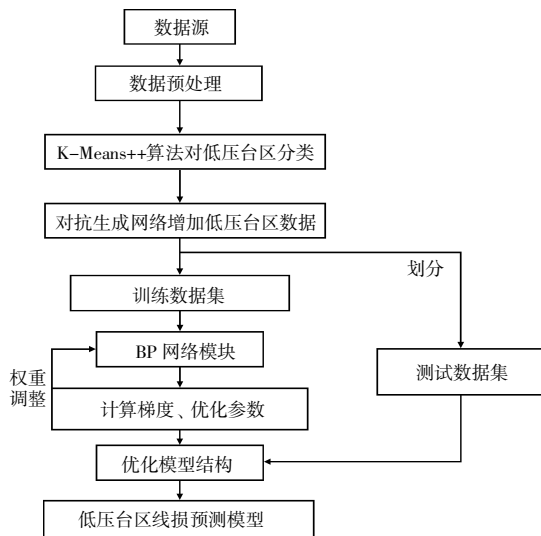


图2 低压台区线损预测模型结构

### 2.1 数据来源及数据集

某市电网公司提供了50 000条低压台区线损相关数据,台区线损率按月统计,空间范围是某市所有低压台区,时间范围是2016年1月至2018年9月。选取的特征属性包括:合计供电量(T\_PPQ)、线损率(L\_LLRL)、上月线损率(L\_LM\_LLRL)、同期线损率(L\_PERIOD\_LLRL)、同期累计线损率(ACCU\_PERIOD\_LLRL)、台区容量(TG\_CAP)、总户数(TOTAL\_USER\_NUM)。

### 2.2 数据处理过程

#### 2.2.1 数据预处理

首先对原始数据进行预处理,包括删除空数据、去掉不符合业务实际的异常数据,再将这些处理后的数据进行归一化处理,最后整理在一张表中保存以备用。归一化处理格式如式(4)所示:

$$\bar{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \quad (4)$$

式中: $x$ 为样本数据, $x=[x_1, x_2, x_3, \dots, x_n]$ , $n$ 为

样本个数; $x_{\max}$ 为样本中的最大值; $x_{\min}$ 为样本中的最小值; $x_i$ 为单个样本。

#### 2.2.2 低压台区数据分类

对于相似的低压台区,其线损率应该是相似的。从保存的表中取出T\_PPQ、TOTAL\_USER\_NUM两个特征值,对所有台区进行K-Means++聚类,计算各个K值对应的轮廓系数,并选择轮廓系数最接近1的聚类数为最优聚类数。各个K值对应的轮廓系数见图3。由图3可知K=2时轮廓系数更接近1。

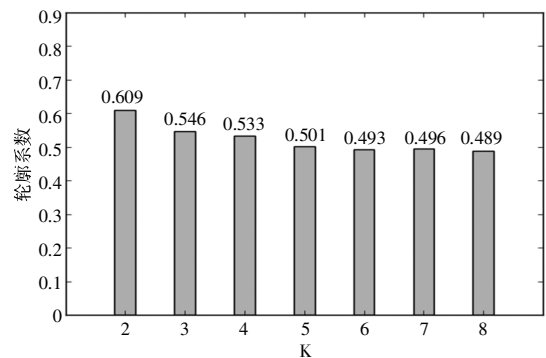


图3 轮廓系数柱状图

通过K-Means++算法把低压台区分为2类,分别记为A类和B类。A类和B类的数据数量见图4。对比A类、B类台区数据发现,A类低压台区大部分是用户数量、供电量均较多的,B类低压台区大部分用户数量、供电量均较少的。

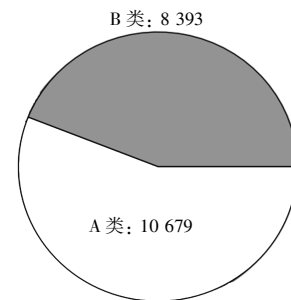


图4 聚类后数量比例饼图

#### 2.2.3 低压台区数据增加

从图4可知A类台区数量为10 679条、B类台区数量为8 393条,对于BP神经网络来说训练的样本数量较少,则可以利用对抗生成网络来训练一个判别网络和一个生成网络,通过生成网络增大样本数据。

真实数据是由 L\_LLRL、L\_LM\_LLRL、L\_PERIOD\_LLRL、ACCU\_PERIOD\_LLRL、TG\_CAP 组成的向量，对抗生成网络噪音数据是(-1, 1)之间的随机数，维度与真实数据相同。将噪音数据输入生成网络里产生伪造数据，判别网络对伪造数据和真实数据进行分辨。生成对抗网络结构如图 5 所示。

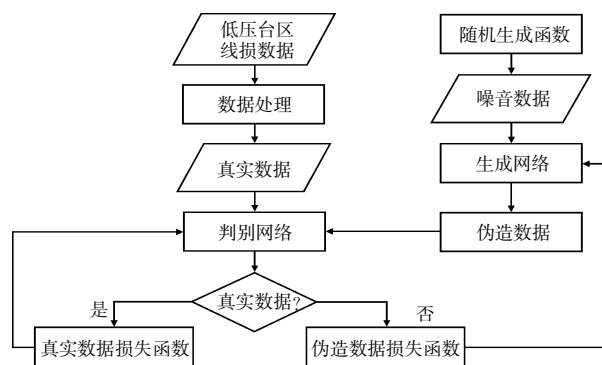


图 5 生成对抗网络结构

随着时间的推移生成网络和判别网络不断进行对抗，使生成网络、判别网络损失函数下降到最小值。A 类、B 类对抗生成网络损失函数分别如图 6、7 所示。从图中可以看出，A 类、B 类对抗生成网络中判别网络均在训练到 300 轮左右时损失函数接近 1，此时获得较好分类效果。在训练到 400 轮时，A 类、B 类对抗生成网络中生成器的参数优化到可以使得损失函数结果较低的程度，说明其可以较为稳定地输出样本数据。

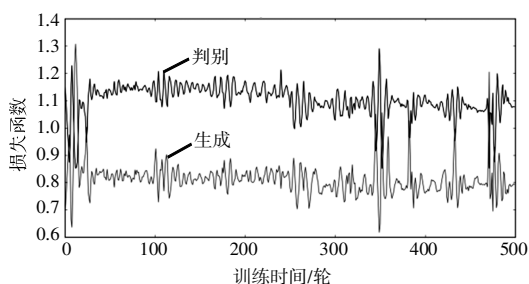


图 6 A 类对抗生成网络损失函数

通过生成器来增加 A 类、B 类低压台区的线损数据，根据实验的实际情况，使用 A 类生成网络、B 类生成网络分别产生 6 000 条线损数据时实验效果最佳。

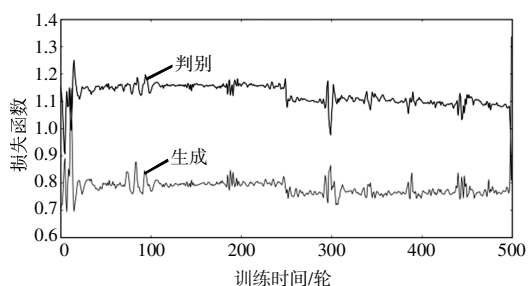


图 7 B 类对抗生成网络损失函数

### 2.3 建立 BP 神经网络的线损率预测模型

BP 神经网络具有结构简单、使用方便、易于理解等特点，函数的逼近能力较强。BP 神经网络的学习过程主要包括信号的前向传播及误差的反向传播两部分。

#### 2.3.1 BP 神经网络超参数选择

BP 神经网络由 3 层全连接的神经网络组成，分别为两层隐藏层和一层输出层。隐藏层神经元个数通常根据经验以及实验尝试获取，本次实验中隐藏层神经元数分别为 64 和 128 个。隐藏层采用 RELU 函数进行激活，增强模型的非线性，初始值选取(-1, 1)之间的随机数。输出层外加一层 softmax 层使神经网络输出变成一个概率分布，最终完成分类预测。经 softmax 回归处理之后的输出为：

$$y'_i = \text{softmax}(y_i) = \frac{e^{y_i}}{\sum_{j=1}^n e^{y_j}} \quad (5)$$

式中： $y_1, y_2, y_3, \dots, y_n$  为原始神经网络的输出； $y'_i$  为输出的概率分布。

学习率直接反映了每次神经网络训练过程中相邻神经元之间连接权值变化速度的快慢，根据实际的训练需要将学习率设置为 0.001，既不会导致函数震荡，也不会延长网络训练的时间。

#### 2.3.2 BP 神经网络的输入数据

数据增加后 A 类、B 类低压台区数据分别为 14 393、16 679 条。把 A 类里的 14 393 个低压台区和 B 类里的 16 679 个低压台区分别拆分为训练集和测试集，拆分比例为 3:1，具体数据数量如表 1 所示。

为 A 类、B 类低压台区数据分别建立预测模型，每个预测模型选取 L\_LM\_LLRL、L\_PERIOD\_LLRL、ACCU\_PERIOD\_LLRL、TG\_CAP 组成向量，

表1 各数据数量

类别	低压台区数量		
	总数量	训练数据	训练数据
A	16 679	12 509	4 170
B	14 393	10 794	3 599

并将此向量数据输入BP神经网络进行训练;选取L<sub>LLR</sub>作为为标签值,根据某电网线损参考数据将线损率分为四类:合格线损率(0<L<sub>LLR</sub><3.5)、欠合格线损率(3.5≤L<sub>LLR</sub><10)、不合格线损率(10≤L<sub>LLR</sub><100)、异常线损率(L<sub>LLR</sub><0, L<sub>LLR</sub>>100),在输入前进行One-Hot编码处理。

### 2.3.3 损失函数与梯度优化

本文使用损失函数作为模型的误差计算方式。通过计算损失函数的值,对模型中的权重进行反馈调节,调整整个模型的参数设置。本次实验所用的损失函数为交叉熵损失函数:

$$H(p, q) = -\sum_x p(x) \log q(x) \quad (6)$$

式中:给定两个概率分布 $p$ 和 $q$ ,通过 $q$ 来表示 $p$ 的交叉熵。

对于上述模型,求导、反馈等操作由Tensorflow后端完成,对权重矩阵进行优化,使梯度逐步下降直到得到最小的损失函数值。

## 2.4 实验结果分析

按本文提出的预测模型分别进行10次低压台区线损预测实验,实验结果为:A类低压台区的平均预测准确率约为92.1%,B类低压台区的平均预测准确率约91.5%。

为了说明对抗生成网络增大的样本数据有效提高预测准确率,在不使用对抗生成网络生成的数据前提下,使用基于BP神经网络的低压台区线损预测模型做了平行对比实验,实验的准确率对比如表2所示。

表2 两种模型的准确率对比

试验方法	预测结果		%
	A类低压台区	B类低压台区	
基于BP神经网络预测模型	89.1	88.4	
本文所提预测模型	92.1	91.5	

## 3 结语

本文提出一种基于对抗生成网络与BP神经网络的低压台区线损预测模型。该模型既充分考

虑了合计供电量、台区容量、总用户等特征数据与当前线损率之间的关系,又使用了上月线损率、同期线损率、同期累计线损率等历史线损数据来提高预测的准确性。所做工作如下:使用Tensorflow框架成功搭建了一个低压台区线损预测模型;某市电网公司提供了某市从2016年1月至2018年9月共50000条低压台区线损相关数据,其中选取了合计供电量、上月线损、台区容量、总用户等特征属性作为特征值,实际中存在大量低压台区缺失特征数据,剔除这些特征数据后还剩下近20000条实验的低压台区线损相关数据。根据相似低压台区线损率相似原则,通过K-Means++算法将低压台区分类,分别记为A类、B类;其数据分别为10679条、8393条;为解决BP神经网络训练数据样本偏少的问题,分类后利用对抗生成网络来增加训练的数据样本,经试验发现A类生成网络、B类生成网络分别产生6000条线损数据时实验结果最佳,再使用增大的样本数据训练BP神经网络从而搭建线损的预测模型。

将有无使用对抗生成网络增大的样本数据做了平行对比实验,实验结果表明,本文提出的基于对抗生成网络与BP神经网络的低压台区线损预测模型具有更加准确的预测效果,表明通过对抗生成网络增大样本数据可以改善低压台区有效线损数据量偏小的问题,从而可以减小不符合业务实际的异常数据、空数据等对线损率研究的影响;快速准确预测低压台区线损率,会帮助供电企业制定合理的降损措施,从而提高企业的经济效益。

## 参考文献:

- [1] 周金华,魏鹏飞,王超.对如何降低线损率的措施浅释[J].中小企业管理与科技(上旬刊),2010(4):236-237.
- [2] WU A, NI B. Line Loss Analysis and Calculation of Electric Power Systems[M]. New York: John Wiley and Sons Ltd, 2015.
- [3] 王海燕.基于核偏最小二乘回归分析的线损率预测[J].计算机仿真,2012,29(11):323-326.
- [4] 郑强.基于灰色理论的电力线损预测[J].中国水运:理论版,2007,5(1):144-145.
- [5] 彭建春,李春晖,祁学红,等.基于快速独立成分分析和支持向量回归的混合馈线线损估算[J].电力系统保护与控制,2012,40(3):51-55.

- [6] 黄韬,刘胜辉,谭艳娜.基于 K-means 聚类算法的研究[J].计算机技术与发展,2011,21(7):54-57.
- [7] ZHANG Y,HEPNER G,DENNISON P.Delineation of phenoregions in geographically diverse regions using K-means++ clustering: a case study in the upper colorado river basin[J].Mapping Sciences & Remote Sensing,2012,49(2):163-181.
- [8] 夏士雄,李文超,周勇,等.一种改进的 K-means 聚类算法[J].东南大学学报(英文版),2007,23(3):435-438.
- [9] 王坤峰,苟超,段艳杰,等.生成式对抗网络 GAN 的研究进展与展望[J].自动化学报,2017,43(3):321-332.
- [10] GOODFELLOW I J,POUGET-ABADIE J,MIRZA M,et al.Generative adversarial networks[J].Advances in Neural Information Processing Systems,2014(3):2672-2680.
- [11] 林懿伦,戴星原,李力,等.人工智能研究的新前线:生成式对抗网络[J].自动化学报,2018,44(5):10-27.
- [12] 余思泉,韩志,唐延东,等.基于对抗生成网络的纹理合成方法[J].红外与激光工程,2018,47(2):43-48.
- [13] JING L,CHENG J H,SHI J Y,et al.Brief Introduction of Back Propagation (BP) Neural Network Algorithm and Its Improvement[J].Advance in Computer Science and Information Engineering,2012,169:553-558.
- [14] 贾丽会,张修如.BP 算法分析与改进[J].计算机技术与发展,2006,16(10):101-103.
- [15] 明东岳,李振华,李俊,等.配网线损同期化计算与多维分析系统[J].电网与清洁能源,2019,35(4):30-36.
- [16] 刘东升,代盛国,商学斌,等.基于压缩感知理论的缺失数据集下线损预测模型[J].广东电力,2019,32(2):80-86.
- [17] 陈亮,吕斌斌,方勤斌,等.基于智能电表海量数据的台区运行态势分析与应用[J].电力大数据,2019,22(2):74-80.
- [18] 张永凯,马欢,张玮,等.计及峰谷分时的配电线路同期月线损预测方法及应用[J].电网与清洁能源,2017,33(10):67-73.
- [19] 王若丞.配电网三相不平衡对线损增加率及电压偏移的影响[J].电力工程技术,2017,36(4):131-136.
- [20] 李健,庄远灿,党三磊,等.基于特征量测量的配电网实时三相线损计算[J].广东电力,2017,30(2):81-86.

---

收稿日期:2019-07-26

作者简介:方舟(1984),男,工程师,研究方向为信息系统安全。

周后盘(1967),男,博士,特聘教授,研究方向为大数据及人工智能应用。(通讯作者)

(本文编辑:李慧敏)