

基于聚类自适应主动学习的电力系统暂态稳定评估

卢东昊¹, 王莉², 张少凡², 蔡燕春², 陈金富¹

(1. 华中科技大学 电气与电子工程学院 强电磁工程与新技术国家重点实验室, 湖北 武汉 430074;

2. 广州供电局有限公司, 广东 广州 510620)

摘要:为了解决电力系统暂态稳定评估中机器学习方法所需样本数多、仿真耗时长的问題,提出主动学习的方法。为了降低主动学习过程中选择样本的冗余度,提出一种聚类自适应主动学习选择策略。通过聚类选择初始样本,使初始样本具有代表性,加快了主动学习进程;将不确定性和代表性2种指标结合自适应地选择权重参数,使选择的样本冗余度低。CEPRI 36节点系统仿真结果表明,主动学习的方法可以在保证准确率的情况下显著降低所需样本数,节省了大量的仿真耗时,参数自适应的选择策略相比传统的不确定性策略大幅提高了训练效率。

关键词:机器学习;主动学习;聚类算法;自适应权重;电力系统;暂态稳定评估

中图分类号:TM 712

文献标志码:A

DOI:10.16081/j.epae.202105024

0 引言

传统的电力系统暂态稳定评估采用时域仿真法,随着电力系统的规模不断扩大,时域仿真法耗时过长的缺点使得其难以应用于快速评估。机器学习的方法因快速准确的特点而被应用于暂态稳定评估中,主要包括决策树DT(Decision Tree)^[1-2]、支持向量机SVM(Support Vector Machine)^[3-5]、人工神经网络ANN(Artificial Neural Network)^[6-7]、极限学习机ELM(Extreme Learning Machine)^[8]、堆叠自动编码器SAE(Stacked AutoEncoder)^[9-11]、卷积神经网络CNN(Convolutional Neural Network)^[12]、深度置信网络DBN(Deep Belief Network)^[13-14]等方法。

文献[2]提出基于二维组合属性决策树的暂态稳定评估方法,与单属性决策树相比具有更好的泛化能力。文献[5]将支持向量机与长短期记忆网络相结合,实现发电机功角轨迹的在线预测。文献[7]利用神经网络构建有功出力控制变量与暂态稳定指标之间的映射关系,可用于暂态稳定预防控制。文献[9]提出一种基于堆叠自动编码器的暂态稳定评估方法来自动提取数据特征,并利用大量无标注样本提高泛化能力,缩短时域仿真时间。文献[12]构建多个相同结构不同参数的卷积神经网络模型进行综合预测,该模型具有良好的性能。文献[14]提出一种基于深度置信网络的暂态稳定评估方法,采用“预训练-微调参数”的方法,在拥有少量训练样本的情况下具有优越的性能。

上述基于机器学习的暂态稳定评估方法大多存

在需要大量带标注样本的问题。文献[9,14]采用无监督学习的思想减少带标注样本的数量,但在如何选择带标注样本上没有进行深入研究。足够的样本数量是保证模型高准确率的前提,样本的生成借助于时域仿真法,因此样本数量决定了生成样本的耗时。在生成的大量样本中难免会存在冗余的样本,带来了额外的仿真成本,因此,人们迫切希望能够缩短仿真时间,同时不损失模型的精度。

本文提出一种基于聚类自适应主动学习CAAL(Clustering Adaptive Active Learning)的电力系统暂态稳定评估模型,使用K-means聚类算法选择初始样本,加快主动学习进程;针对主动学习过程中每轮迭代可能选出冗余样本的问题,提出一种将不确定性和代表性相结合的综合指标,并采用自适应的方法动态调整不确定性和代表性2个指标的权重,提高主动学习的性能。在CEPRI 36节点测试系统上进行仿真,结果表明本文提出的方法能够显著减少仿真样本数,缩短仿真时间,且该方法性能优于传统主动学习方法。

1 主动学习概述

在机器学习领域中,根据训练集中样本是否带标注可以将机器学习分成监督学习、无监督学习和半监督学习。传统的监督学习需要大量的带标注样本才能达到较高的精度,但为了得到大量的带标注样本存在一些问题^[15]:在某些领域样本的总量很少,获取新样本的成本高;在已有大量样本中带标注的样本少,大量样本未被标注,监督学习很难达到很高的训练精度;样本中存在低质量样本和冗余样本,将这些样本加入训练集反而可能降低模型的鲁棒性。无监督学习是一种针对含有大量未标注样本问题的方法,它可以不需要样本的标注,挖掘样本集自身的

收稿日期:2020-08-09;修回日期:2021-03-16

基金项目:广州供电局有限公司科技项目(080016KK52170007)

Project supported by the Science and Technology Project of Guangzhou Power Supply Bureau Co.,Ltd.(080016KK52170007)

特性进行学习,但其性能通常略逊于监督学习。

如何在大量的未标注样本中选取部分样本进行标注,再将其加入训练集进行学习,从而提升分类器的性能,是机器学习的研究热点之一。近年来,主动学习在机器学习领域受到广泛关注,其在图像分类、工业故障识别、模式类别挖掘等领域应用较多,而在电力领域应用较少。文献[16]将主动学习引入暂态稳定评估,采用基于熵值不确定性的选择策略,但没有考虑样本的代表性以及初始样本集的构建问题。

主动学习的思想是从大量未标注的样本中寻找质量高的样本进行人工标注,再将其加入训练集中进行学习。由于标注是人工进行的,因此不会引入错误标注,且能显著降低标注成本以及提升分类器的性能。同时,根据设计的选择策略,可以防止选择的样本与已有训练样本冗余度过高,进一步提升训练样本集的质量。

主动学习模型可以用 $A=(C, L, U, S, Q)$ 表示。其中, A 为主动学习模型; C 为分类器; L 和 U 分别为已标注样本集和未标注样本集; S 为能标注样本的专家或机器; Q 为选择策略。根据采样方法的不同,主动学习的选择策略大体可以分为基于流的采样方法和基于池的采样方法。基于流的采样方法从未标注样本池中将样本依次输入 S , 若满足指定的条件则进行标注, 否则直接舍弃该样本。由于基于流的采样方法需要预先设定一个标注条件, 而不同的任务往往需要不同的条件, 因此该方法很难成为一种通用方法。基于池的采样方法按照预设的规则从未标注样本池中选择样本交给分类器 C , 当分类器 C 难以对其分类时进行人工标注。与基于流的采样方法相比, 该方法每次都可以选出对分类贡献度最高的样本, 降低了查询和标注的成本, 因此被广泛使用。基于池的采样方法包括不确定性标准、泛化误差缩减标准等。本文在不确定性标准的基础上增加代表性指标, 并将不确定性与代表性指标自适应结合, 使得筛选的样本对分类器准确率有更高的贡献度。

2 CAAL

本文提出 CAAL 的方法, 利用聚类的方法选择初始样本, 将不确定性与代表性结合, 以解决主动学习中样本信息冗余的问题。

2.1 初始样本的选择

主动学习需要一组初始样本来构建基准分类器, 一般采用随机方法选择初始样本。初始样本的好坏决定了基准分类器的性能, 并影响后续主动学习的进程, 而随机挑选的初始样本通常不具有代表性。如果能够构建一组具有代表性的初始样本集, 则能提高基准分类器的性能, 加快主动学习进程。

电力系统运行方式数据的不同维度间存在物理

规律的时空耦合, 且存在复杂的非线性相关性, 通过随机运行方式生成的样本具有一定的分布规律。在样本未标注的情况下, 通过聚类的方法可得知样本的大体分布, 离聚类中心越近的点越能代表当前类别, 样本越具有代表性。基于此, 本文采用 K -means 聚类算法对所有未标注样本进行聚类后挑选初始样本, 使得挑选出的样本具有代表性。具体方法为: 所有样本均未标注, 利用 K -means 算法将所有样本聚为 k 类, 由算法根据聚类效果挑选 k 的具体取值, 选择离 k 个聚类中心最近的若干样本进行标注, 将这部分标注样本作为初始样本训练基准分类器。

2.2 不确定性指标

基于不确定性的选择策略是主动学习中最常用的一种方法, 该方法选择当前模型中最不确定的样本进行标注。常用的不确定性指标有概率不确定性、熵值不确定性和距离不确定性。本文采用样本的信息熵来衡量样本的不确定性, 熵值越大, 当前分类器越不能确定样本所属的类别。样本的熵值定义为:

$$f(x_i) = - \sum_{y \in Y} P(y|x_i, \theta_L) \ln P(y|x_i, \theta_L) \quad (1)$$

其中, Y 为所有可能的分类取值集合, 本文中电力系统暂态稳定评估为二分类问题, 故 Y 元素为 0 和 1; θ_L 为在当前已标注样本集上训练得到的分类器; $P(y|x_i, \theta_L)$ 为当前分类器对未标注样本 x_i 预测为 0 或 1 的概率值。

基于熵值不确定性的选择策略选择信息熵最大的样本:

$$i^* = \operatorname{argmax}_{x_i \in U} f(x_i) \quad (2)$$

其中, argmax 函数的值为 $f(x_i)$ 取最大值时对应的索引值; i^* 为每轮迭代选择的样本的编号。

2.3 代表性指标

在主动学习问题中, 使用基于不确定性的选择策略已经可以达到较好的效果, 但不确定性只反映样本信息量的大小, 若只考虑样本的不确定性, 每轮迭代选择的样本可能与已标注样本之间冗余度较高, 影响主动学习的效果。样本的代表性对模型性能也有一定影响。

以图 1 所示的二分类问题为例进行说明。图中实心图形表示已标注样本, 空心图形表示未标注样本, 圆形和方形分别表示二分类中的 2 类, 实线表示分类超平面。

由图 1 可知: 未标注样本 1 离超平面的距离比未标注样本 2 更近, 样本 1 的不确定性更大, 但在样本 1 的附近有 2 个距离其很近的已标注样本, 可以认为这 3 个样本的相似度很高, 此时若将样本 1 加入训练集中, 对分类器的性能提升作用不大; 样本 2 的不确定性不如样本 1, 但样本 2 附近的已标注样本离其较

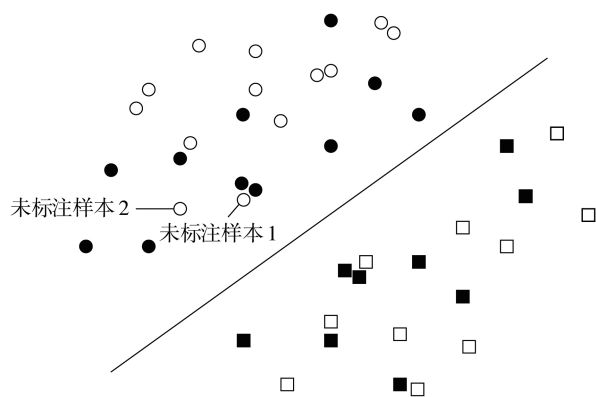


图1 主动学习不确定性和代表性权衡

Fig.1 Trade-off between uncertainty and representativeness in active learning

远,其代表性高于样本1,将其加入训练集后可以降低训练集的冗余度,其对分类器的提升作用可能高于样本1。因此,本文将不确定性和代表性相结合,提出一种参数自适应的选择策略,以改善主动学习中训练样本冗余的问题。

本文采用样本间的欧氏距离衡量样本的代表性。在每轮迭代中,计算当前已标注样本与未标注样本间的欧氏距离,距离越小则说明样本离得越近,样本的相似度越高。2个样本间的欧氏距离定义为:

$$d_{\text{euclidean}}(X, Y) = \sqrt{\sum_{i=1}^n (x'_i - y'_i)^2} \quad (3)$$

其中, $X = (x'_1, x'_2, \dots, x'_n)$ 、 $Y = (y'_1, y'_2, \dots, y'_n)$ 为2个样本向量; n 为样本的特征维数。

对某个未标注样本 x_i , 用其与当前所有已标注样本间欧氏距离的最小值 $d(x_i)$ 衡量样本的代表性:

$$d(x_i) = \min_{x_j \in U, x_j \in L} d_{\text{euclidean}}(x_i, x_j) \quad (4)$$

$d(x_i)$ 越大,说明样本离当前已标注样本越远,样本越具有代表性。

2.4 参数自适应

为了降低样本的冗余度,将不确定性和代表性相结合,选择不确定性高同时具有代表性的样本,本文采用加权的形式将2个指标结合,得到综合指标 $g_\beta(x_i)$ 为:

$$g_\beta(x_i) = \beta f(x_i) + (1 - \beta) d(x_i) \quad (5)$$

其中, $\beta \in [0, 1]$ 为权重系数。当 $0.5 < \beta \leq 1$ 时,不确定性指标占的权重更大,特别地,当 $\beta = 1$ 时,综合指标等价于不确定性指标;当 $\beta = 0.5$ 时,不确定性指标与代表性指标的权重相同;当 $0 \leq \beta < 0.5$ 时,综合指标更偏向于代表性指标。然而,如何确定 β 的取值是个困难的问题,且在每轮迭代中分类器的参数一直变化,固定的 β 值很难保证在变化的过程中选择的样本是最佳的。为此,本文设置 β 的取值集合为

$\{0, 0.1, \dots, 1\}$, 当 β 取这11个不同的值时可以选出11个不同的样本,由于对样本进行仿真标注耗时较长,故先对11个样本的标注结果进行假设。分别将这11个样本中的每个样本标注为0和1,再分别重新训练模型,比较这11个样本中的每个样本对当前模型带来的影响。如果假设的标注与样本的真实标注相反,则重新训练模型会导致模型准确率下降;如果假设的标注与样本的真实标注相同,则可以提高模型准确率。分别计算新的模型在测试集上的准确率,将所有准确率由高到低排序,将准确率最高的样本作为本轮迭代最终选择的样本,交由仿真软件标注后将其加入已标注样本池,并将其从未标注样本池中移除。虽然采用枚举法选择权重系数 β 会导致计算量的增加,但是可以减少所需标注样本的数量以及标注样本的计算量,枚举法带来的增加的计算量占比较小,整体计算量是一个下降的趋势。

3 基于CAAL的暂态稳定评估模型

3.1 算法流程

基于CAAL的暂态稳定评估算法步骤如下。

(1) 获取运行数据。通过数据采集与监视控制SCADA(Supervisory Control And Data Acquisition)系统状态估计可获取大量未标注的稳态运行数据,若未包含某些运行方式下的数据,则可通过离线仿真进行补充,将这些数据作为初始未标注样本池。

(2) 选择初始样本。对所有未标注样本进行K-means聚类,选取聚类中心附近样本,模拟三相接地短路故障,由仿真软件计算暂态稳定性,对样本进行标注,将其作为初始已标注样本池,并将其从未标注样本池中移除,在已标注样本池上训练基准分类器。

(3) 主动学习。根据不同的 β 值计算未标注样本池中每个样本的综合指标,选出一组子样本集,再计算子样本集中每个样本加入已标注样本池后模型的准确率,选择准确率最高的样本,仿真计算暂态稳定性,对样本标注后将其加入已标注样本池,并将其从未标注样本池中移除。在更新后的已标注样本池上重新训练分类器。

(4) 终止条件。当主动学习达到预设的迭代次数或者分类器达到预设精度要求时,停止训练,否则重复步骤(3)。

最终的分器可用于暂态稳定性预测,该方法可得到预测精度高的分类器,且所需标注的样本数量少,可节省大量离线仿真时间,提高模型训练效率。

3.2 模型的输入和输出

模型的输入特征选为系统的稳态特征量,可通过SCADA系统状态估计获得,包括各节点的电压幅值、相角,发电机的有功出力、无功出力,负荷的有功

功率、无功功率以及各线路的有功功率、无功功率。模型的输出为故障下的暂态稳定类别,包括稳定和失稳。稳定判据通过计算暂态稳定指数 I_{TSI} 获得:

$$I_{TSI} = \frac{360^\circ - |\Delta\delta_{\max}|}{360^\circ + |\Delta\delta_{\max}|} \quad (6)$$

其中, $\Delta\delta_{\max}$ 为任意 2 台发电机间的最大相对功角差。 $I_{TSI} > 0$ 时判定为稳定, $I_{TSI} \leq 0$ 时判定为失稳。

3.3 性能评价

电力系统暂态稳定评估是一个非平衡分类问题,稳定样本一般多于失稳样本,且漏判的代价比误判的代价高。为此,除了分类准确率外,引入非平衡问题中应用广泛的 F1 值 P_{F1} 。F1 值的计算依赖于暂态稳定评估结果的混淆矩阵,如表 1 所示。其中, I_{TP} 和 I_{FP} 分别为预测为稳定的样本中真实结果为稳定和失稳的数目; I_{FN} 和 I_{TN} 分别为预测为失稳的样本中真实结果为稳定和失稳的数目。

表 1 暂态稳定评估混淆矩阵

Table 1 Confusion matrix of transient stability assessment

预测结果	混淆矩阵	
	真实结果为稳定	真实结果为失稳
稳定	I_{TP}	I_{FP}
失稳	I_{FN}	I_{TN}

模型的分类准确率 P_{Acc} 为预测结果与真实结果相同的样本比例,即:

$$P_{Acc} = \frac{I_{TP} + I_{TN}}{I_{TP} + I_{FP} + I_{FN} + I_{TN}} \quad (7)$$

精确率 $P_{Precision}$ 为模型预测为失稳的样本中预测正确的比例。召回率 P_{Recall} 为真实结果为失稳的样本中预测正确的比例。精确率和召回率有时会出现矛盾的情况, F1 值综合考虑这 2 个指标, F1 值越高,模型的性能越好。

$$P_{Precision} = \frac{I_{TN}}{I_{FN} + I_{TN}} \quad (8)$$

$$P_{Recall} = \frac{I_{TN}}{I_{TN} + I_{FP}} \quad (9)$$

$$P_{F1} = \frac{2P_{Precision}P_{Recall}}{P_{Precision} + P_{Recall}} \quad (10)$$

在参数自适应过程中,需计算不同样本加入训练集后准确率的变化,选取加入后准确率增加最大的样本进行标注。在每轮迭代后,计算当前分类器的分类准确率和 F1 值,并绘制其随迭代次数增加的变化曲线。

用所有样本训练模型,得到的准确率为该模型在当前样本集上的上限,设置主动学习的终止条件为达到该上限,根据准确率达到上限时所需样本数量来衡量主动学习算法的性能。

4 算例分析

4.1 数据集构造

算例采用 CEPRI 36 节点系统,该系统包含 36 个节点、8 台发电机和 26 条输电线路,系统单线图见附录中图 A1。

利用电力系统分析综合程序 PSASP 生成数据集,考虑负荷水平在基准负荷的 60%~140% 之间随机取值,生成不同的潮流方式。经过筛选,选出 2 条发生故障后易导致暂态失稳的线路 AC_{10} 和 AC_{11} ,见附录中图 A1。在线路上设置三相接地短路故障,故障位置位于线路全长的 20% 处,故障持续时间设置为 0.2 s,仿真时长为 3 s。最终各获得 2 214 个样本,其中故障 1 包含 1 684 个稳定样本和 530 个失稳样本,故障 2 包含 1 658 个稳定样本和 556 个失稳样本。选取其中的 20% 为测试样本,剩余样本作为训练样本,并将其设为未标注样本,作为初始未标注样本池。

4.2 性能评估

本文使用的分类器是支持向量机,采用径向基函数,通过五折交叉验证和网格搜索寻找支持向量机的最优参数,最终取惩罚系数为 100,径向基函数参数为 0.01。在选取训练样本和测试样本时,对样本进行随机划分,在不同的训练集上进行学习,并将 5 次的计算结果进行平均。将本文提出的 CAAL 与随机选择(Random)、基于熵值不确定性策略(Unc)和未进行初始聚类的自适应主动学习方法(AAL)相比较。首先比较相同迭代次数时各方法的准确率,设置初始已标注样本的数量为 50 个,每轮迭代选择 1 个样本进行标注,迭代次数分别为 50 次和 100 次。故障 1 和故障 2 的 50 次和 100 次迭代时准确率随迭代次数的变化曲线分别如图 2 和图 3 所示。

图 2 和图 3 中的虚线为用所有训练样本学习后模型的准确率,其为模型在该组样本上准确率的上限。由图中可见,CAAL 在迭代开始时准确率高于其他方法,这说明经过聚类选择出的初始样本具有一定的代表性,比随机选择初始样本的方法更好,而 2 种采用自适应主动学习的方法(CAAL 和 AAL)在迭代过程中选择的样本兼具不确定性和代表性,比

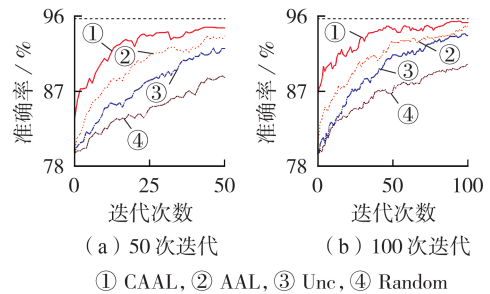


图 2 故障 1 的分类准确率曲线

Fig.2 Classification accuracy rate curves of Fault 1

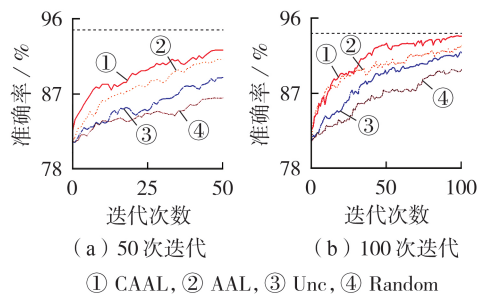


图3 故障2的分类准确率曲线

Fig.3 Classification accuracy rate curves of Fault 2

Random和Unc准确率更高。表2和表3分别为故障1和故障2的50次和100次迭代后各方法的准确率、F1值、误判率(真实结果为稳定,预测为失稳)和漏判率(真实结果为失稳,预测为稳定),以及达到准确率上限时所需的样本数量。

表2 故障1不同方法性能比较

Table 2 Performance comparison among different methods for Fault 1

迭代次数	参数	CAAL	AAL	Unc	Random
50	准确率 / %	94.49	93.54	92.01	88.53
	F1值	0.9644	0.9568	0.9468	0.9207
	误判率 / %	8.68	9.58	13.87	14.14
	漏判率 / %	4.41	5.50	5.99	9.84
100	准确率 / %	95.08	94.71	93.63	90.20
	F1值	0.9661	0.9615	0.9580	0.9341
	误判率 / %	5.26	6.52	9.73	11.40
	漏判率 / %	3.33	3.60	4.38	8.09
达到准确率上限所需样本数		194	265	288	955

表3 故障2不同方法性能比较

Table 3 Performance comparison among different methods for Fault 2

迭代次数	参数	CAAL	AAL	Unc	Random
50	准确率 / %	92.19	91.06	88.94	86.50
	F1值	0.8891	0.8774	0.8546	0.8274
	误判率 / %	8.56	9.06	9.34	12.85
	漏判率 / %	7.36	9.25	11.38	12.37
100	准确率 / %	93.77	92.69	91.92	89.80
	F1值	0.9197	0.9065	0.8929	0.8725
	误判率 / %	6.98	7.20	7.99	10.15
	漏判率 / %	6.75	8.89	9.15	11.35
达到准确率上限所需样本数		319	405	442	1062

从表2和表3中可以看出,本文所提CAAL方法在相同迭代次数时的准确率和F1值均高于其他方法,且样本的误判率和漏判率均低于其他方法,达到准确率上限所需样本数最少,在保证准确率的前提下可显著降低标注代价。与Random方法相比,本文所提方法所需带标注样本仅为其所需样本数的25%左右,可节省大量时域仿真标注样本的时间。与其他主动学习方法相比,本文所提方法的性能也更优。

5 结论

本文引入主动学习的思想,提出一种基于CAAL的电力系统暂态稳定评估方法,并在CEPRI 36节点系统上进行仿真验证。仿真结果表明:

(1)与传统的机器学习方法相比,主动学习能够在保证模型准确率的基础上,显著降低标注代价,缩短标注样本的时间,提高训练效率;

(2)利用K-means聚类算法选择初始样本,能够提高初始样本的代表性,为主动学习提供了良好的初值,加快了主动学习进程;

(3)主动学习的选择策略将不确定性和代表性相结合,改进了传统主动学习方法仅考虑不确定性的不足,使得每轮迭代选择的样本对模型性能的提升作用更大;

(4)通过自适应的方式改变不确定性和代表性的权重,确保每轮迭代过程选择的样本最优。

本文初步探索了主动学习在暂态稳定评估中的应用,为如何选择训练样本提供了一种新思路。由于主动学习在每次迭代时需重新训练模型,本文选择浅层机器学习方法以加快训练效率。如何将主动学习与深度学习相结合,并充分利用未标注样本将主动学习与半监督学习或无监督学习相结合,以提高模型准确率和训练效率,将是接下来研究的重点。另外,通常同一样本中的不同特征对暂态稳定性的影响不同,而且在不同样本中的同一特征对暂态稳定性的影响也不尽相同,如何更加准确地衡量样本的代表性值得进一步讨论与研究。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] GENC I, DIAO R S, VITTAL V, et al. Decision tree-based preventive and corrective control applications for dynamic security enhancement in power systems[J]. IEEE Transactions on Power Systems, 2010, 25(3): 1611-1619.
- [2] 王康, 孙宏斌, 张伯明, 等. 基于二维组合属性决策树的暂态稳定评估[J]. 中国电机工程学报, 2009, 29(增刊1): 17-24. WANG Kang, SUN Hongbin, ZHANG Boming, et al. Transient stability assessment based on 2D combined attribute decision tree[J]. Proceedings of the CSEE, 2009, 29(Supplement 1): 17-24.
- [3] 戴远航, 陈磊, 张玮灵, 等. 基于多支持向量机综合的电力系统暂态稳定评估[J]. 中国电机工程学报, 2016, 36(5): 1173-1180. DAI Yuanhang, CHEN Lei, ZHANG Weiling, et al. Power system transient stability assessment based on multi-support vector machines[J]. Proceedings of the CSEE, 2016, 36(5): 1173-1180.
- [4] WANG B, FANG B W, WANG Y J, et al. Power system transient stability assessment based on big data and the core vector machine[J]. IEEE Transactions on Smart Grid, 2016, 7(5): 2561-2570.
- [5] 刘俐, 李勇, 曹一家, 等. 基于支持向量机和长短期记忆网络的暂态功角稳定预测方法[J]. 电力自动化设备, 2020, 40(2): 129-139.

- LIU Li, LI Yong, CAO Yijia, et al. Transient rotor angle stability prediction method based on SVM and LSTM network[J]. Electric Power Automation Equipment, 2020, 40(2): 129-139.
- [6] 姚德全, 贾宏杰, 赵帅. 基于复合神经网络的电力系统暂态稳定评估和裕度预测[J]. 电力系统自动化, 2013, 37(20): 41-46.
- YAO Dequan, JIA Hongjie, ZHAO Shuai. Power system transient stability assessment and stability margin prediction based on compound neural network[J]. Automation of Electric Power Systems, 2013, 37(20): 41-46.
- [7] 杨跃, 刘友波, 刘俊勇, 等. 基于神经网络预测校核的暂态稳定预防控制[J]. 电网技术, 2018, 42(12): 4076-4084.
- YANG Yue, LIU Youbo, LIU Junyong, et al. Preventive transient stability control based on neural network security predictor[J]. Power System Technology, 2018, 42(12): 4076-4084.
- [8] 陈振, 肖先勇, 李长松, 等. 基于代价敏感极端学习机的电力系统暂态稳定评估方法[J]. 电力自动化设备, 2016, 36(2): 118-123.
- CHEN Zhen, XIAO Xianyong, LI Changsong, et al. Power system transient stability assessment based on cost-sensitive extreme learning machine[J]. Electric Power Automation Equipment, 2016, 36(2): 118-123.
- [9] 朱乔木, 陈金富, 李弘毅, 等. 基于堆叠自动编码器的电力系统暂态稳定评估[J]. 中国电机工程学报, 2018, 38(10): 2937-2946, 3144.
- ZHU Qiaomu, CHEN Jinfu, LI Hongyi, et al. Transient stability assessment based on stacked autoencoder[J]. Proceedings of the CSEE, 2018, 38(10): 2937-2946, 3144.
- [10] 王怀远, 陈启凡. 基于堆叠变分自动编码器的电力系统暂态稳定评估方法[J]. 电力自动化设备, 2019, 39(12): 134-139.
- WANG Huaiyuan, CHEN Qifan. Transient stability assessment method of electric power systems based on stacked variational auto-encoder[J]. Electric Power Automation Equipment, 2019, 39(12): 134-139.
- [11] 尹雪燕, 闫炯程, 刘玉田, 等. 基于深度学习的暂态稳定评估与严重度分级[J]. 电力自动化设备, 2018, 38(5): 64-69.
- YIN Xueyan, YAN Jiongcheng, LIU Yutian, et al. Deep learning based transient stability assessment and severity grading[J]. Electric Power Automation Equipment, 2018, 38(5): 64-69.
- [12] 田芳, 周孝信, 史东宇, 等. 基于卷积神经网络综合模型和稳态特征量的电力系统暂态稳定评估[J]. 中国电机工程学报, 2019, 39(14): 4025-4032.
- TIAN Fang, ZHOU Xiaoxin, SHI Dongyu, et al. Power system transient stability assessment based on comprehensive convolutional neural network model and steady-state features[J]. Proceedings of the CSEE, 2019, 39(14): 4025-4032.
- [13] 胡伟, 郑乐, 闵勇, 等. 基于深度学习的电力系统故障后暂态稳定评估研究[J]. 电网技术, 2017, 41(10): 3140-3146.
- HU Wei, ZHENG Le, MIN Yong, et al. Research on power system transient stability assessment based on deep learning of big data technique[J]. Power System Technology, 2017, 41(10): 3140-3146.
- [14] 朱乔木, 党杰, 陈金富, 等. 基于深度置信网络的电力系统暂态稳定评估方法[J]. 中国电机工程学报, 2018, 38(3): 735-743.
- ZHU Qiaomu, DANG Jie, CHEN Jinfu, et al. A method for power system transient stability assessment based on deep belief networks[J]. Proceedings of the CSEE, 2018, 38(3): 735-743.
- [15] 杨文柱, 田潇潇, 王思乐, 等. 主动学习算法研究进展[J]. 河北大学学报(自然科学版), 2017, 37(2): 216-224.
- YANG Wenzhu, TIAN Xiaoxiao, WANG Sile, et al. Recent advances in active learning algorithms[J]. Journal of Hebei University(Natural Science Edition), 2017, 37(2): 216-224.
- [16] 谢培元, 袁文, 刘永刚, 等. 基于主动学习的电力系统暂态稳定评估方法[J/OL]. 电测与仪表. [2019-11-26]. <http://kns.cnki.net/kcms/detail/23.1202.th.20191115.1208.017.html>.

作者简介:



卢东昊

卢东昊(1996—),男,浙江东阳人,硕士研究生,主要研究方向为机器学习与人工智能技术在电力系统稳定性中的应用(E-mail: M201871354@hust.edu.cn);

陈金富(1972—),男,福建莆田人,副教授,博士,通信作者,主要研究方向为电力系统分析计算与运行控制、继电保护整定计算等(E-mail: chenjinfu@mail.hust.edu.cn)。

(编辑 王锦秀)

Transient stability assessment of power system based on clustering adaptive active learning

LU Donghao¹, WANG Li², ZHANG Shaofan², CAI Yanchun², CHEN Jinfu¹

(1. State Key Laboratory of Advanced Electromagnetic Engineering and Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China;

2. Guangzhou Power Supply Bureau Co., Ltd., Guangzhou 510620, China)

Abstract: In order to solve the problems that large number of samples and long simulation time are needed by machine learning method in transient stability assessment of power system, an active learning method is proposed. In order to reduce the redundancy of sample selection in active learning progress, a clustering adaptive active learning selection strategy is proposed. The initial samples are selected by clustering, which makes the initial samples representative and speeds up the active learning process. The two indicators of uncertainty and representativeness are combined to adaptively select the weight parameters, which makes the redundancy of the selected samples low. The simulative results of CEPRI 36-bus system show that the active learning method can significantly reduce the needed number of samples while ensuring the accuracy, which saves much simulation time, and the adaptive parameter selection strategy greatly improves the training efficiency compared with the traditional uncertainty strategy.

Key words: machine learning; active learning; clustering algorithms; adaptive weight; electric power systems; transient stability assessment

附录：

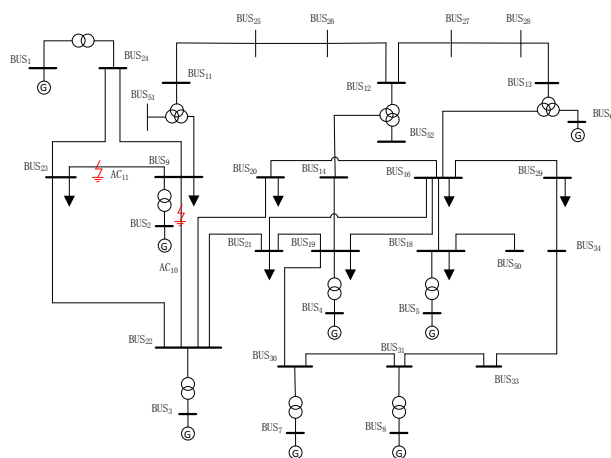


图 A1 CEPRI 36 节点系统单线图

Fig.A1 Single-line diagram of CEPRI 36-bus system