

基于XGBoost-EE的电力系统暂态稳定评估方法

邬春明^{1,2},任继红²

(1. 东北电力大学 现代电力系统仿真控制与绿色电能新技术教育部重点实验室,吉林 吉林 132012;
2. 东北电力大学 电气工程学院,吉林 吉林 132012)

摘要:深度学习在暂态稳定评估中发挥着越来越重要的作用,然而电网规模逐渐扩大导致数据出现维数灾难,这对模型的性能提出了更高的要求。目前,暂态稳定特征构建需要依靠人工经验,具有主观性;深度学习的模型在设计和训练上耗时、耗力。针对以上两点,结合极限梯度提升(XGBoost)算法和实体嵌入(EE)网络,提出了一种基于XGBoost-EE的电力系统暂态稳定评估方法。首先通过XGBoost算法的路径规则生成类别特征,将原始特征进行降维。然后采用EE网络对新的特征进行分类,从而完成快速、精准的暂态稳定评估。该方法充分利用了机器学习算法处理速度快和神经网络评估精度高的优点,能够直接面向底层量测数据。最后,在IEEE新英格兰10机39节点和IEEE 50机145节点系统的仿真结果表明,所提方法相比于其他方法具有更高的预测精度和更好的抗噪性能,且在训练时不容易过拟合。

关键词:XGBoost算法;实体嵌入;暂态稳定评估;深度学习;大数据

中图分类号:TM 712

文献标志码:A

DOI:10.16081/j.epae.202011032

0 引言

随着经济的不断发展,电网互联的规模不断加大,其运行环境也越来越复杂。如果系统在遭受大扰动后工作人员没有及时地采取相应的安全稳定控制措施,那么将会导致电压失稳或线路解列等现象,甚至还有可能导致大面积停电等严重事故。因此需要探索鲁棒性高且快速准确的暂态稳定评估TSA(Transient Stability Assessment)方法,以便运行人员在遇到不安全系统状态时能够及时采取预防控制措施。在传统暂态稳定评估方法中,时域仿真法^[1]评估准确高,但计算花费时间较长;直接法^[2-4]能快速提供暂态稳定评估结果,但只能给出近似的结果,并且很难应用于复杂的电力系统。

近年来,学者们开始研究基于人工智能方法的电力系统暂态稳定评估,通过大量训练样本学习系统特征与稳定结果之间的复杂非线性关系,模型训练完成后在实际使用时可以快速对电力系统进行评估。建模时间仅仅耗费在样本积累和模型训练这种线下环节,而在预测环节消耗的时间很少。文献[5]提出了一种将支持向量机SVM(Support Vector Machine)与长短期记忆LSTM(Long Short-Term Memory)网络相结合的预测方法,实现了对暂态功角稳定性及功角轨迹的预测。文献[6]提出了基于多属性决策树DT(Decision Tree)和电网暂态稳定裕度指标的电网稳定规则辨识方法,利用大量数值仿真进行样本训练和DT分析,获取了电网稳定关联规则。文献[7]使用人工神经网络ANN(Artificial Neural Network)进行暂态稳定分析。文献[8]将改进的极限梯

度提升(XGBoost)算法用于暂态稳定评估,在输出部分引入Logistic函数将输出概率化,既满足了二分类的需求,又能够在一定程度上判断XGBoost模型预测的可靠程度。

由于浅层学习模型有一定的局限性,当特征维数多时,模型的表征能力有限,计算复杂度高,泛化能力差。因此,深度学习作为机器学习研究中的一个新领域,被广泛应用于电力系统暂态稳定评估。文献[9]提出了一种安全域概念下的基于堆叠降噪自动编码器SDAE(Stacked Denoised AutoEncoder)特征集成模型的暂态稳定评估和严重度分级方法,使用SDAE提取的层次特征和SVM模型分别作为集成学习模型的输入和子学习器,基于效用理论对运行方式进行严重度分级,使暂态稳定评估结果更加直观。文献[10]首次对深度学习模型的每一层转换后的特征使用t-分布随机邻近嵌入t-SNE(t-distributed Stochastic Neighbor Embedding)进行降维可视化。文献[11]针对电力系统暂态稳定评估中稳定样本与不稳定样本误分类代价不同的特点,提出了一种基于代价敏感极限学习机ELM(Extreme Learning Machine)的电力系统暂态稳定评估方法。此外,在深度学习中常用于图像分类、自然语言处理以及其他数据分析任务的卷积神经网络CNN(Convolutional Neural Network)^[12]等方法也被应用到暂态稳定评估领域。现有的基于人工智能方法的暂态稳定评估模型均获得了较高的评估准确率,但仍存在一些不足:第一,深度学习的模型一般较为复杂,数据之间的非线性函数关系全面,容易过拟合,并且大部分模型训练时间较长,模型离线更新成本较高;第二,基于深度学习的暂态稳定评估模型大多无法

可视化,被认为是“黑箱模型”,不具有可解释性,很难应用于实际。

本文针对现阶段电力系统暂态评估方法的不足,结合 XGBoost 算法和实体嵌入(EE)网络,提出了一种基于 XGBoost-EE 的电力系统暂态稳定评估方法。所提方法使用 XGBoost 算法对原始特征进行快速的编码,编码后的特征矩阵维数较少,能够有效克服电力数据的维数灾难问题,加速实体嵌入网络的训练;实体嵌入将特征自身属性以及特征间关系映射到欧几里得空间,使编码后的特征空间能够进行可视化分析。本文采用新英格兰 10 机 39 节点系统和 IEEE 50 机 145 节点系统进行仿真分析,并将本文模型与其他模型进行对比,运用 t-SNE^[13]降维方法对 XGBoost 转换后的特征以及实体嵌入关系图进行了可视化展示。

1 算法介绍

1.1 XGBoost 算法

XGBoost 算法是基于梯度提升 DT 的改进算法。它可以有效地构建增强树并能够并行运行,树的集成模型如下:

$$\hat{y}_i = \sum_{k=1}^K f_k(\mathbf{x}_i) \quad f_k \in F \quad (1)$$

其中, K 为树的数目; F 为树的集合空间; \hat{y}_i 为模型的预测值; \mathbf{x}_i 为第 i 个数据点的特征向量,且 $\mathbf{x}_i \in \mathbf{R}^M$ ($i=1,2,\dots,n$; n 为给定样本数目); f_k 为第 k 棵独立树的结构,与叶子的权重向量 w 有关。

XGBoost 分类器由学习模型、参数调整和优化目标函数组成。学习模型是基础函数和权重相结合的结果;参数调整是构建模型和优化模型的必要过程;目标函数的优化程度决定了模型的准确率。目标函数的优化程度越高,模型的泛化能力就越强。因此,通过最小化损失函数和增加模型复杂度的惩罚项来实现对目标函数的优化。XGBoost 模型的目标损失函数由两部分组成:

$$L = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2)$$

其中,第一部分 $\sum_{i=1}^n l(y_i, \hat{y}_i)$ 为真实值 y_i 和预测值 \hat{y}_i 之间的误差函数;第二部分 $\sum_{k=1}^K \Omega(f_k)$ 为树的复杂度函数,用于控制模型复杂度的正则项。 $\Omega(f_k)$ 可以表示为:

$$\Omega(f) = \gamma T + 0.5\lambda \|w\|^2 \quad (3)$$

其中, γ 和 λ 为模型的惩罚因子; T 为树上的叶子数。式(3)在最小化序列的过程中,在训练的每一轮增加增量函数 $f_i(\mathbf{x}_i)$ 。因此,目标函数可以改写为:

$$L_{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_i(\mathbf{x}_i)) + \Omega(f_i) \quad (4)$$

其中, t 表示第 t 轮迭代。对于式(4),使用二阶泰勒级数展开式将第 j 棵树的每片叶子中的样本集合定义为 $I_j = \{i \mid q(\mathbf{x}_i = j)\}$,其中 q 为树结构。目标损失函数的一阶导数是 $g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$,二阶导数是 $h_i = \partial_{\hat{y}_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)})$,由此可得:

$$L_{(t)} \cong \sum_{i=1}^n \left(g_i f_i(\mathbf{x}_i) + 0.5 h_i f_i^2(\mathbf{x}_i) \right) + \Omega(f_i) \cong \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) w_j + 0.5 \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \quad (5)$$

定义 $G_j = \sum_{i \in I_j} g_i$ 、 $H_j = \sum_{i \in I_j} h_i$,式(5)可简化为:

$$L_{(t)} \cong \sum_{j=1}^T \left[G_j w_j + 0.5 (H_j + \lambda) w_j^2 \right] + \gamma T \quad (6)$$

其中, w_j 的偏导数为:

$$w_j' = -\frac{G_j}{H_j + \lambda} \quad (7)$$

将权重向量 w 代入目标损失函数得:

$$L_{(t)} \cong -0.5 \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T \quad (8)$$

由式(8)可知,目标损失函数越小,模型越好。

1.2 实体嵌入

从机器学习的角度出发,电力系统暂态稳定性评估是一个高维分类问题。解决这个复杂问题的关键是找到输入变量 e_i 和输出变量 y 的非线性映射关系,可表示为:

$$y = f(e_1, e_2, \dots, e_n) \quad (9)$$

为了使模型更好地学习到式(9)的近似值,将结构化数据中的每个特征映射到一个向量,如下:

$$\alpha_i: e_i \rightarrow e_i \quad (10)$$

特征到向量的映射等价于独热编码层之上连接了额外的线性神经元,如图 1 所示。为了说明这一点,将 e_i 的独热编码表示为式(11)。

$$\beta_i: e_i \rightarrow \delta_{e,\alpha} \quad (11)$$

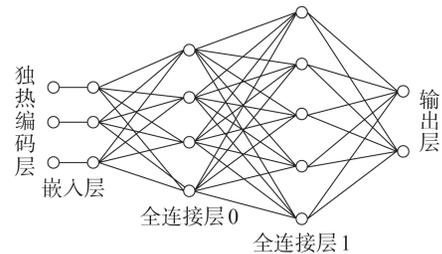


图 1 嵌入层可以理解在独热编码层之上建立的额外层
Fig.1 Embedding layer can be understood as extra layer above one-hot coding layer

其中, $\delta_{e_i, \alpha}$ 为一个克罗内克函数, $\alpha=e_i$ 时, 元素 δ 非零且唯一, α 为独热编码层的索引值。

如果输入变量 e_i 的数量为 m , 则 $\delta_{e_i, \alpha}$ 对应于长度为 m 的向量。当给定输入 e_i 时, 其额外的线性神经元的输出可以定义如下:

$$x_i = \sum_{\alpha} \eta_{\alpha\beta} \delta_{e_i, \alpha} = \eta_{e_i, \beta} \quad (12)$$

其中, $\eta_{\alpha\beta}$ 为独热层与嵌入层之间的连接权重; β 为嵌入层的索引。显然, 嵌入的映射是 2 个连接层的权重, 它类似于神经网络中的权重参数, 可以通过训练进行优化。

将分类变量嵌入实体中, 并将它们组合成一个层用作神经网络的普通输入层, 其他层可以在其之上进行构建, 整个网络可以用反向传播方法进行训练。通过这种方式, 实体嵌入层可以了解每种类别的内在属性, 而更深的层则形成它们的复杂组合。嵌入尺寸是超参数, 通常在 $1 \sim m$ 或 $1 \sim m-1$ 之间选择。在仿真中, 设置此超参数有以下 2 个准则: ①粗略估计每个类别可以描述多少个不同的有意义的值, 并将其用作嵌入维度; ②如果不能满足准则①, 那么从 $1 \sim m-1$ 中选择嵌入的尺寸, 对于嵌入的尺寸需要更多的数学理论指导。

通过实体嵌入的方法将分类变量的相似值映射到更近距离的欧几里得空间。文献[14]证明了将原始度量空间等距嵌入欧几里得空间的必要和充分条件。如果实体嵌入是等距嵌入, 应用式(9), 需要满足式(13)所示的矩阵是一个正定矩阵, 具体可参考文献[14]。

$$(M_i)_{pq} = e^{-\lambda \langle f(e_i^p, \bar{e}_i) - f(e_i^q, \bar{e}_i) \rangle_{e_i}} \quad (13)$$

通过计算可得, 原始潮流数据经 XGBoost 转换后的特征矩阵是正定的。由式(13)可知, 本文使用的实体嵌入是等距嵌入。

2 暂态稳定评估模型

2.1 XGBoost-EE

XGBoost 算法是实现非线性到元组转换的非常简单而强大的算法[15]。其将每棵树视为一个类别特征, XGBoost 模型输出的值是样本落在该棵树的叶子索引值, 以获得转换后的特征矩阵。

神经网络非常适合处理连续数据, 它在计算机视觉、语音识别和自然语言处理领域取得了良好的效果, 但它在处理结构化数据时优势并不那么突出[16]。当引入实体嵌入时, 神经网络可达到良好的效果。实体嵌入增加了特征连续性, 以一种有意义的方式将网络限制在较小的空间内, 加快了神经网络的收敛。更确切地说, 实体嵌入方法迫使神经网络学习每个特征的内在属性及其在空间上的分布。

本文所提模型的整体结构如图 2 所示。将从相

量测量单元直接得到的原始潮流数据经过预处理之后, 导入模型。首先通过 XGBoost 算法将原始特征转换成类别特征, 然后将新的特征用作嵌入层的输入, 最后输出到全连接层进行分类。模型有明显的加速效果, 原因如下: ①XGBoost 算法在设置树的数量 (N_{round}) 参数时并不需要如文献[17]选取较多的树, 这对于算法有加快的作用; ②转换后的特征矩阵特征维度仅为 $N_{\text{round}} + 1$, 相比于原始特征维数大幅减小, 使实体嵌入网络训练加快。

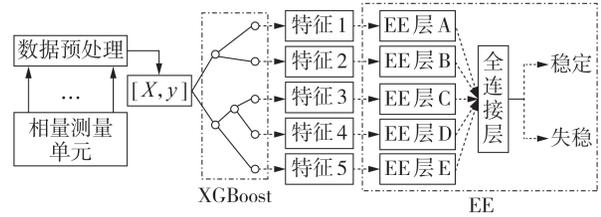


图2 模型整体结构

Fig.2 Whole structure of model

2.2 输入特征的选择

本文提出的评估模型旨在从底层量测数据中获取评估知识, 因此底层数据须较为全面地反映系统的动态行为。电力系统的暂态稳定性直接取决于发电机转子角度的动态特性, 因此通常采用转子角度作为输入特征[18]。另外, 发电机有功功率和无功功率对发电机的转子角度也有影响, 因此本文选取发电机转子角度和有功功率、无功功率作为输入特征。输入的样本特征集 X 可以表示为以下矩阵形式:

$$\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,v} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,v} \\ \vdots & \vdots & & \vdots \\ x_{u,1} & x_{u,2} & \cdots & x_{u,v} \end{bmatrix} \quad (14)$$

其中, u 为输入样本的数量; v 为特征的尺寸。

2.3 模型的输出

模型的最后一层具有 2 个输出节点, 分别对应于稳定和不稳定类别。稳定的判据参照暂态稳定指数 I_{TSl} :

$$I_{\text{TSl}} = \frac{360^\circ - |\Delta\delta_{\max}|}{360^\circ + |\Delta\delta_{\max}|} \quad (15)$$

其中, $\Delta\delta_{\max}$ 为任意 2 台发电机之间的最大功角之差。当 $I_{\text{TSl}} < 0$ 时, 系统不稳定; 当 $I_{\text{TSl}} > 0$ 时, 系统稳定。

2.4 评价指标

为了更好地评估模型的性能, 本文选择总体准确率 P_{Acc} 和接受者操作特征 ROC (Receiver Operating Characteristic) 曲线下坐标轴所包围的面积 (AUC) 值 P_{AUC} 这 2 个指标, 对模型进行综合评估。 P_{Acc} 为:

$$P_{\text{Acc}} = \frac{I_{\text{TP}} + I_{\text{TN}}}{I_{\text{TP}} + I_{\text{TN}} + I_{\text{FP}} + I_{\text{FN}}} \quad (16)$$

ROC 曲线的 x 轴为假正类率 P_{FPR} , y 轴为真正类率 P_{TPR} , 分别见式(17)和式(18)。

$$P_{FPR} = \frac{I_{FN}}{I_{TN} + I_{FN}} \quad (17)$$

$$P_{TPR} = \frac{I_{TN}}{I_{TN} + I_{FP}} \quad (18)$$

其中, I_{TP} 为正确分类的稳定样本的数量; I_{TN} 为正确分类的不稳定样本的数量; I_{FN} 为错误分类的不稳定样本的数量; I_{FP} 为错误分类的稳定样本的数量。 P_{Acc} 和 P_{AUC} 的取值是范围 $[0, 1]$ 中的实数。在评估模型性能时, 需要综合这 2 个指标。

3 算例分析

3.1 数据集生成

本文所使用的编程语言为 Python, 深度学习框架选择基于 TensorFlow 的 Keras 框架, 计算机配置为 Intel core i5-6200 CPU 8 GB。仿真系统采用 IEEE 新英格兰 10 机 39 节点系统和 IEEE 50 机 145 节点系统, 仿真平台为 PSD-BPA。算例考虑了 80%、85%、...、130% 共计 11 个负荷水平, 并相应地改变发电机出力以保证系统功率平衡且各母线电压维持在 0.95~1.05 p.u. 的合理范围之内。在输电线路设置故障, 故障位置位于线路的 0、20%、50%、80% 处, 故障类型为三相永久性短路故障, 故障持续时间分别设为 0.1、0.15、0.18、0.2 s, 仿真时长设为 5 s。对于 IEEE 39 节点系统共生成 5 984 个样本, 其中失稳样本 1 932 个, 稳定样本 4 052 个; 对于 IEEE 145 节点系统共生成 29 520 个样本, 其中失稳样本 5 536 个, 稳定样本 23 984 个。

3.2 模型参数及性能评估

本文使用 XGBoost 算法进行特征转换, 并对转换后的特征进行实体嵌入, 在嵌入层之上使用 2 层全连接层进行判稳。在 IEEE 39 节点系统中, XGBoost-EE 参数设置为: 树的最大深度为 8; 学习率为 0.2; 树的数量为 50。嵌入层的嵌入尺寸设置为每列的类别数, 即每棵树包含叶节点的数目。隐含层神经单元的个数采用遍历法确定, 遍历区间为 $[50, 300]$, 经搜寻, 隐含层单元个数设定为 $\{150, 50\}$, 输出层为 2 个神经元。在 IEEE 145 节点系统中, XGBoost-EE 参数设置为: 树的最大深度为 7; 学习率为 0.4; 树的数量为 52。嵌入层的设置同 IEEE 39 节点系统, 隐含层神经单元的个数最终设定为 $\{500, 300\}$ 。

将本文提出的方法与 CNN、堆叠自编码器 SAE (Stacked AutoEncoder)、ANN、XGBoost、逻辑回归 LR (Logistics Regression)、SVM、随机森林 RF (Random Forest)、DT 评估算法进行比较。为了确保得到最佳效果, 在使用其他分类器时使用了网格搜索方法来优化参数。

在 IEEE 39 节点系统中, CNN 采用 {双层卷积+Relu 激活层, 池化, Dropout 层, 全连接输出} 结构的神经网络结构, 卷积核大小为 3×3 , 池化层窗口大小为 2×2 , Dropout 层系数设为 0.25, 全连接层为 1 层; SAE 采用 5 层深度网络结构, 每层神经元数逐层减少, 隐含层单元数目为 300-200-150-80-30; ANN 的网络结构与 SAE 相同, 隐含层单元数目为 200-300-150-100-50。SVM 采用径向基核函数 RBF (Radial Basis Function), 惩罚因子 C 的取值范围为 $\{0.1, 1, 10, 100, 1000, 10000\}$, 核参数 φ 的取值范围为 $\{0.5, 1, 2, 4, 8, 16, 32\}$, 最终得到 $C=100, \varphi=4$; RF 的基分类器数量为 100, 最大深度为 10; DT 采用 C4.5 算法, 最大深度为 9。考虑到各种评估方法的随机性, 实验采用了 5 折交叉验证, 评估准确性是交叉验证的平均值, 结果如表 1 所示。

表 1 不同暂态稳定评估模型在 IEEE 39 节点系统的测试性能比较

Table 1 Comparison of test performance of different TSA models in IEEE 39-bus system

模型	P_{AUC}	$P_{Acc} / \%$	训练时间 / s
LR	0.9659	96.60	16.40
SVM	0.9978	96.10	6.47
RF	0.9732	97.28	7.04
DT	0.9511	94.92	47.83
XGBoost	0.9646	97.33	0.5
SAE	0.9842	98.11	380.0
ANN	0.9451	91.78	414.7
CNN	0.9899	98.33	514.5
XGBoost-EE	0.9934	98.96	28.5

从表 1 中可以看出, XGBoost-EE 模型的平均评估精度比 LR、SVM、RF、DT、XGBoost、SAE、ANN、CNN 分别高了 2.81%、2.225%、2.98%、2.31%、1.56%、1.095%、6.62%、0.755%。另外, 浅层模型 (如 LR、SVM、RF、DT、XGBoost) 虽然在小系统上有较好的表现, 但当系统拓扑规模扩大时, 评估表现远远不如深层模型 (如 SAE、CNN), 而本文模型在拓扑结构复杂时仍能保持良好的精度和较快的计算速率。综上所述, 本文模型融合了机器学习与深度学习处理数据的优点, 具有较好的泛化能力和较高的准确性, 表明了使用 XGBoost 构建新特征非常适合实体嵌入, 可以显著提高算法性能。

在 IEEE 145 节点系统中, CNN 采用 4 层卷积+池化层和全连接输出的神经网络结构, 卷积层和池化窗口设置同 39 节点系统, 全连接层为 2 层; SAE 的隐含层神经元设定为 300-400-500-300-200-50 共 6 层; ANN 的网络结构与 SAE 相同, 隐含层单元数目为 200-300-500-150-100-50; SVM 中 $C=100, \varphi=4$ 。RF 的基分类器数量为 120, 最大深度为 18; DT 最大深度为 12。对比结果如附录中表 A1 所示。

4 模型的可视化

4.1 XGBoost 转换特征可视化

使用可视化工具 t-SNE 将 2 套 IEEE 测试系统的 20 维人工特征数据^[17]和 XGBoost 构建的特征数据映射到三维空间以进行比较,结果如图 3 所示,子图中左侧为人工特征降维图,右侧为 XGBoost 构建的特征降维图。从图中可以看出,在人工特征空间中的 2 类样本掺杂,难以区分;而在 XGBoost 转化后的特征空间中, XGBoost 构建的特征表达几乎可以将稳定的样本与不稳定的样本分离,仅存在少量的样本重叠。更重要的是,人工特征提取可能造成大量有价值的信息丢失^[19],而且只能表征电力系统的整体运行状况,平等地对待了所有特征信息,不能准确地表征系统发生故障的运行状况。

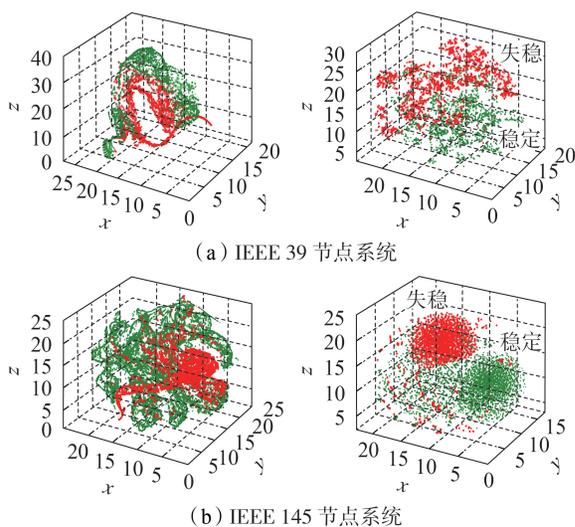


图3 人工特征与 XGBoost 算法构建特征的可视化对比图
Fig.3 Visual comparison diagram of artificial features and XGBoost algorithm constructed features

4.2 实体嵌入可视化

如 1.2 节所述,本文所使用的实体嵌入为等距嵌入,将 XGBoost 转换后的新特征映射到欧几里得空间。由于新特征的每一列中编号相同的叶节点都被嵌入欧几里得空间中的相同位置,如图 4 所示。为了便于说明实体嵌入的作用,本文选取 IEEE 39 节点系统中母线电压作为输入特征对模型进行可视化分析,随机选取数据集下训练的 XGBoost 模型中的第 44 棵树进行可视化。通过 t-SNE 方法降维将嵌入矩阵映射到二维空间中,以观察各个叶节点的嵌入位置。附录中图 A1 展示了 XGBoost 的第 44 棵树的提取规则,对其 9 个叶节点进行编号,这 9 个叶节点在树生成的新特征中也代表 9 个唯一值。经统计发现,在图 A1 中的 9 个不同叶节点中,落在白色节点上的大多数是稳定样本,而落在黑色节点上的大多数是不稳定样本。这些特点在图 4 中是相同的,图中

的节点颜色越接近黑色,落到该节点的不稳定样本数目就越多;节点颜色越接近白色,落到该节点的稳定样本数目越多。

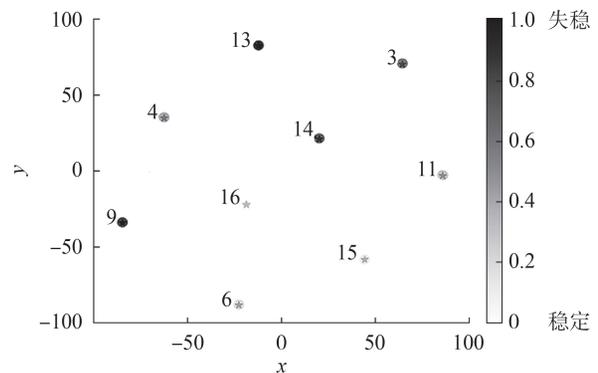


图4 XGBoost 第 44 棵树的实体嵌入图

Fig.4 Entity embedded diagram of XGBoost 44th tree

实体嵌入提供了一种分析 XGBoost 转换后特征的方法,即从嵌入位置和距离分析叶节点之间的关系,这有助于提取特征间更深层次的规则。可以看到,具有相似颜色的样本被嵌入相近的欧几里得空间中,并且每个叶节点的位置都是唯一的。不同树的嵌入规则并不相同,例如在图 4 中具有更多稳定样本的叶节点主要集中在右下部分。

5 模型抗噪性能分析

考虑到实际系统中的测量工况易受较多因素干扰,根据 IEEE C37.118 标准模拟相量测量单元的量测误差^[20],分别在两系统的原始数据集中添加不同信噪比的高斯白噪声,实验结果如附录中图 A2 所示。

加入噪声后的模型中,除 XGBoost-EE 之外的其他 5 种评估模型都受到了比较明显的影响,随着噪声水平的增加,准确率 P_{Acc} 和 P_{AUC} 大幅下降。由于 XGBoost 转换后的特征是类别特征,具有离散性,因此噪声给模型带来的影响较小,评估指标并未出现跌落式下降,当噪声水平达到 10 dB 时准确率 P_{Acc} 和 P_{AUC} 仅分别下降了 10.13% 和 11.02%。而其他模型较容易受到干扰,当噪声水平达到 10 dB 时出现了跌落式下降。

6 结论

本文针对电力系统暂态稳定评估问题,提出了一种基于 XGBoost-EE 的电力系统暂态稳定评估方法。在 IEEE 新英格兰 10 机 39 节点系统和 IEEE 50 机 145 节点系统上进行了仿真研究,得到如下结论:

(1) 将 XGBoost 算法用于特征转换,对特征进行了离散化处理,改善了模型的抗噪性能;

(2) 所提方法能够自适应多种数据集,借助深层模型自动提取数据特征完成分类评估,比 LR、SVM、

XGBoost、ANN、SAE 等常规评估方法拥有更高的评估准确率和更高的 P_{AUC} ;

(3) 采用的实体嵌入为等距嵌入,在空间位置上有唯一对应的关系,能够有效防止网络的过拟合,从而加快收敛速度。

本文将深度学习中的实体嵌入与电力系统暂态稳定评估进行结合,为该课题的研究提供了新思路。由于本文采用的实体嵌入是等距嵌入,这为研究更深层次的规则提供了可能。另外,对有限度量空间中实体嵌入关系的深入理解还需要进一步的探究。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] ZADKHA S, JATSKEVICH J, VAAHEDI E. A multi-decomposition approach for accelerated time-domain simulation of transient stability problems[J]. IEEE Transactions on Power Systems, 2015, 30(5): 2301-2311.
- [2] CHANG H D, CHU C C, CAULEY G. Direct stability analysis of electric power systems using energy functions: theory, applications and perspective[J]. Proceedings of the IEEE, 1995, 83(11): 1497-1529.
- [3] XUE Y, VAN CUTSEM T, RIBBENS-PAVELLA M. A simple direct method for fast transient stability assessment of large power systems[J]. IEEE Transactions on Power Systems, 1988, 3(2): 400-412.
- [4] 刘笙. 电力系统暂态稳定分析的能量函数方法[J]. 电网技术, 1995, 19(2): 11-17.
LIU Sheng. Transient energy function analysis for power system stability[J]. Power System Technology, 1995, 19(2): 11-17.
- [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] 石访, 张林林, 胡熊伟, 等. 基于多属性决策树的电网暂态稳定规则提取方法[J]. 电工技术学报, 2019, 34(11): 2364-2374.
SHI Fang, ZHANG Linlin, HU Xiongwei, et al. Power system transient stability rules extraction based on multi-attribute decision tree[J]. Transactions of China Electrotechnical Society, 2019, 34(11): 2364-2374.
- [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] 张晨宇, 王慧芳, 叶晓君. 基于 XGBoost 算法的电力系统暂态稳定评估[J]. 电力自动化设备, 2019, 39(3): 77-83, 89.
ZHANG Chenyu, WANG Hui芳, YE Xiaojun. Transient stability assessment of power system based on XGBoost algorithm[J]. Electric Power Automation Equipment, 2019, 39(3): 77-83, 89.
- [9] 尹雪燕, 同炯程, 刘玉田, 等. 基于深度学习的暂态稳定评估与严重程度分级[J]. 电力自动化设备, 2018, 38(5): 64-69.
YIN Xueyan, TONG Jiongcheng, LIU Yutian, et al. Deep learning based transient stability assessment and severity grading[J]. Electric Power Automation Equipment, 2018, 38(5): 64-69.
- [10] 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.
- [11] 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.
- [12] ZHOU Y, GUO Q, SUN H, et al. A novel data-driven approach for transient stability prediction of power systems considering the operational variability[J]. International Journal of Electrical Power & Energy Systems, 2019, 107: 379-394.
- [13] GISBRECHT A, SCHULZ A, HAMMER B. Parametric nonlinear dimensionality reduction using kernel t-SNE[J]. Neurocomputing, 2015, 147(1): 71-82.
- [14] SCHOENBERG I J. Metric spaces and positive definite functions[J]. Transactions of the American Mathematical Society, 1938, 44(3): 522-536.
- [15] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM International Conference on Knowledge Discovery and Data Mining. San Francisco, USA: [s.n.], 2016: 85-97.
- [16] GUO C, BERKHAHN F. Entity embedding of categorical variables [DB/OL]. [2019-12-26]. <https://arxiv.org/abs/1604.06737>.
- [17] 陈明华, 刘群英, 张家枢, 等. 基于 XGBoost 的电力系统暂态稳定预测方法研究[J]. 电网技术, 2020, 44(3): 1026-1034.
CHEN Minghua, LIU Qunying, ZHANG Jiashu, et al. XGBoost-based algorithm for post-fault transient stability status prediction[J]. Power System Technology, 2020, 44(3): 1026-1034.
- [18] GUO Tingyan, MILANOVIC J V. Probabilistic framework for assessing the accuracy of data mining tool for online prediction of transient stability[J]. IEEE Transactions on Power Systems, 2014, 29(1): 377-385.
- [19] 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.
- [20] HE M, VITTAL V, ZHANG J. Online dynamic security assessment with missing PMU measurements: a data mining approach[J]. IEEE Transactions on Power Systems, 2013, 28(2): 1969-1977.

作者简介:



邬春明

邬春明(1966—),男,吉林吉林人,教授,主要研究方向为电力系统暂态分析(E-mail: wuchunming@neepu.edu.cn);

任继红(1996—),女,吉林吉林人,硕士研究生,通信作者,主要研究方向为人工智能在电力系统暂态分析中的应用(E-mail: 1491235676@qq.com)。

(编辑 李玮)

(下转第 152 页 continued on page 152)

Comprehensive assessment of power system dynamic stability based on continuous wavelet transform

LIU Chunxiao¹,JIANG Tao²,LI Xue²,LI Peng¹,ZHANG Jianxin¹,LI Zhiyong¹,YANG Huanhuan¹

(1. Power Dispatching and Control Center,China Southern Power Grid Co.,Ltd.,Guangzhou 510623,China;

2. School of Electrical Engineering,Northeast Electric Power University,Jilin 132012,China)

Abstract: A comprehensive assessment method of dominant oscillation modes, oscillation mode shapes, participation factors and coherent generator groups for power system is proposed based on CWT (Continuous Wavelet Transform). Firstly, the time-frequency domain decomposition of multi-channel wide area measurement information in power system is realized by using wavelet transform, and the wavelet coefficient matrix corresponding to each measurement channel is obtained. For each wavelet coefficient matrix, the key wavelet coefficient vectors strongly related to the dominant oscillation modes of system are identified by wavelet power spectrum. Then, the oscillation frequency and damping ratio of dominant oscillation modes are assessed based on the key wavelet coefficient vectors. On this basis, the system's oscillation mode under each dominant oscillation mode are assessed by means of the cross wavelet transform. Furthermore, according to the obtained oscillation modes, the participation factor and coherent generator group of each measurement channel are assessed, and then the comprehensive assessment of power system dynamic stability is realized based on CWT. The proposed method is used to analyze the simulation data of 16-machine 68-bus test system and the wide area measured data of an actual power grid, and the results verify the effectiveness and feasibility of the proposed method.

Key words: wide area measurement information; continuous wavelet transform; dominant oscillation modes; oscillation mode shapes; participation factors; coherent generator groups; dynamic stability; electric power systems

(上接第143页 continued from page 143)

Power system transient stability assessment method based on XGBoost-EE

WU Chunming^{1,2},REN Jihong²

(1. Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology,

Ministry of Education,Northeast Electric Power University,Jilin 132012,China;

2. School of Electrical Engineering,Northeast Electric Power University,Jilin 132012,China)

Abstract: Deep learning plays an increasingly important role in transient stability evaluation. However, the increase of power system scale generally results in dimension disasters. In this case, an efficient and tractable computation model is highly desirable. Currently, the construction of transient stability features generally relies on the experience of power system operators, which is more or less subjective. However, the deep learning approach is generally time-consuming and labor-intensive in aspects of design and training. Based on the above two points, a transient stability assessment method of power system based on XGBoost-EE is developed by combining XGBoost (eXtreme Gradient Boosting) algorithm and EE (Entity Embedding) network. Firstly, the path rules of the tree are extracted and the category features are generated by XGBoost algorithm. In this way, the original features are dimensionally reduced. Then, the EE network is used to classify the new features, which provides a fast and accurate assessment. The proposed method, hence, takes full advantage of the fast processing speed of machine learning algorithms and the high accuracy of neural network evaluation. Simulative results based on IEEE New England 10-machine 39-bus system and IEEE 50-machine 145-bus system show that the proposed method exhibits higher prediction accuracy and better anti-noise performance than other approaches. Additionally, the proposed method is not easy to become over-fit during the training process.

Key words: XGBoost algorithm; entity embedding; transient stability assessment; deep learning; big data

附录

表 A1 不同暂态稳定评估模型在 IEEE 145 节点系统测试性能比较

Table A1 Comparison of test performance of different TSA models in IEEE 39-bus system

模型名称	P_{AUC}	$P_{Acc}/\%$	训练时间/s
LR	0.9299	96.47	122.72
SVM	0.9994	98.14	494.07
RF	0.9948	97.61	181.62
DT	0.9911	99.15	1579
XGBoost	0.9946	98.04	0.7
SAE	0.9850	98.39	471.5
ANN	0.9415	93.67	697.8
CNN	0.9708	98.85	956.4
XGBoost-EE	0.9993	99.73	60.7

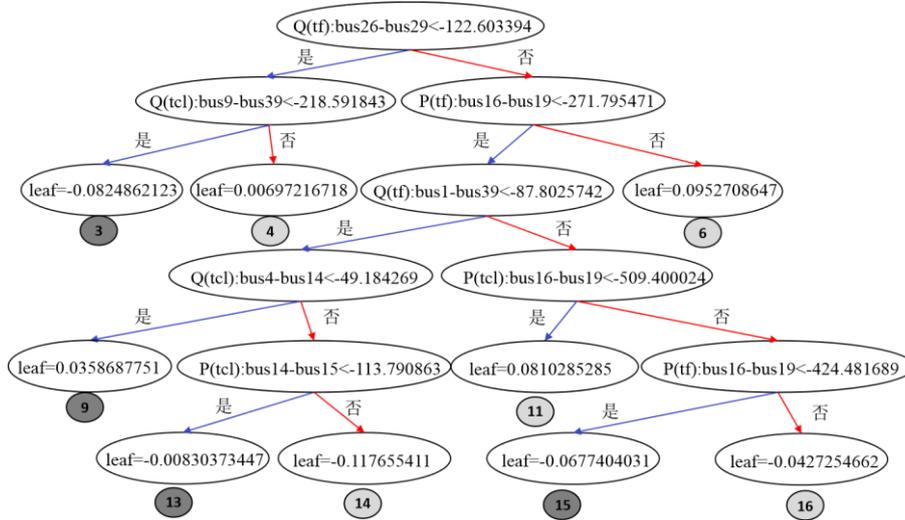
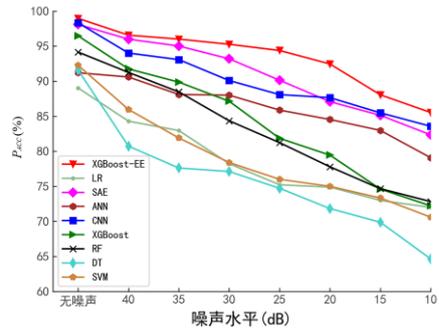
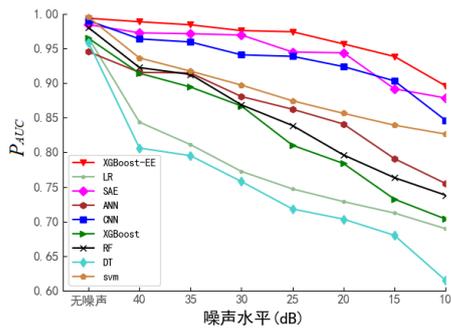
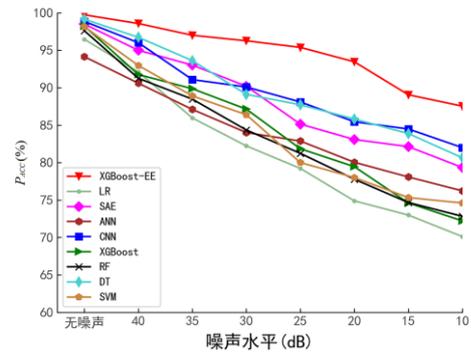
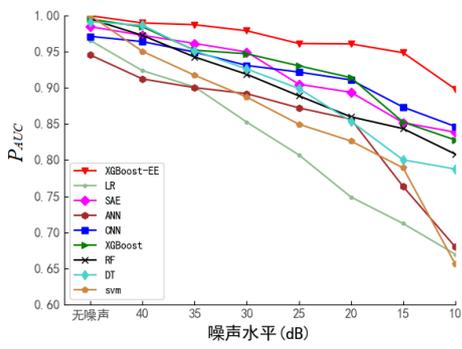


图 A1 XGBoost 的第 44 棵树

Fig.A1 XGBoost 44th tree



(a) 10机39节点系统



(b) 50机145节点系统

图 A2 抗噪能力测试图

Fig.A2 Test diagram of anti-noise ability