Power System Transient Stability Assessment Based on Stacked Autoencoders and Support Vector Machine

Jingxuan Tang, Huibin Sui*
School of Electrical Engineering, Shandong University, Shandong, China.

*Corresponding author e-mail: tjxaihg@163.com

Abstract. Transient stability assessment is the key point to ensure the safe and stable operation of the power system. Based on the data collected from the phasor measurement unit (PMU), a transient stability assessment method combining stacked automatic encoder (SAE) and support vector machine (SVM) is proposed. Multi-layer abstract learning is performed on the original features by the SAE, and the extracted feature is used to train the SVM and test. Simulation on the New England 39-bus test system show that the proposed model has high accuracy.

1. Introduction
With the continuous operation of UHV AC and DC projects, the interconnection between large-scale power grids and the continuous access of new energy and electric vehicles, the operating point of the power system is getting closer to the stability limit. Its security and stability are facing a severe test. Online evaluation of the security and stability of large power grids based on data mining technology provides a new idea for intelligent scheduling.

After suffering from large disturbances such as short-circuit faults, cut-off loads, etc., it is necessary to quickly assess whether the power system is stable. Traditional transient stability assessment methods include time domain simulation, direct method, and extended area method. However, these methods do not fully meet the needs of calculation accuracy, speed and capacity, and are difficult to apply to online calculations. At the same time, with the large deployment of PMUs, data mining methods with high precision, low dimensionality and meeting the requirements of online evaluation speed are applied to transient stability assessment, such as artificial neural networks, support vector machines, decision trees and extreme learning machines.

This paper introduces the deep learning model, combines the feature extraction ability of the stacked automatic encoder and the classification performance of the support vector machine. By simulating on the New England 39-bus test system, compares and analyzes the numerical results of the simulation to verify the superiority of the proposed assessment model.

2. Stacked automatic encoder
SAE is a typical deep neural network, and its basic building block is an autoencoder (AE). AE is a symmetric neural network consisting of an input layer, a hidden layer and an output layer [1]. Its structure is shown in Figure 1.
Figure 1. Autoencoder structure

For a training data set \( \{x_1, x_2, \ldots, x_n\} \) of capacity \( n \), the encoding and decoding processes are as shown in equations (1) and (2), respectively. Where \( \omega_1 \) and \( \omega_2 \) are respectively an encoding matrix and a decoding matrix, and \( b_1 \) and \( b_2 \) are respectively an encoding offset vector and a decoding offset vector, 
\[
\lambda = \frac{1}{1 + \exp(-x)}.
\]

\[
h = \lambda (\omega_1 x + b_1) \tag{1}
\]

\[
\hat{x} = \lambda (\omega_2 h + b_2) \tag{2}
\]

The training process is to find the optimal parameters \( \omega \) and \( b \), so that the reconstruction error is small enough to make the features of the original data included in the coding matrix. Among them,
\[
L(x, \hat{x}) = \sum_{i=1}^{n} \|x_i - \hat{x}_i\|
\]

SAE is a deep neural network model which stacking of multiple AE. The output of the lower layer AE will be used as the input to the upper layer AE. The gradual abstraction of features is achieved by stacking of AEs, ultimately resulting in more compact representative data features.

3. Support vector machine

SVM is a machine learning method based on statistical learning theory that put forward by Vapnik et al. Its main idea is to find an optimal hyperplane in high dimensional space as the segmentation surface of the sample, and to maximize the classification interval.

The description of SVM classification problem is: when a set of training sets \( S = \{(x_i, y_i) | x_i \in R^n, y_i \in R\}_{i=1}^{l} \) is given, solve the optimal classification hyperplane. Where \( x_i \) is the input feature of the \( i \)-th description system, \( y_i \) is the classifier identifier of the \( i \)-th sample, and \( l \) is the sample number. If the training set is linearly separable, the optimal classification function \( f(x) = \text{sgn}(w \cdot x + b) \) is determined by solving the following optimization problems.
In the formula, \( C > 0 \) is the penalty factor, and \( \xi_i \) is the slack variable. Introduce the Lagrange multipliers \( \alpha_i, \gamma_i \) and construct the Lagrange function as follows:

\[
L(w, b, \xi, \alpha, \gamma) = \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i - \sum_{i=1}^{l} \alpha_i \left( y_i [(w^T \cdot x_i) + b] - 1 + \xi_i \right) - \sum_{i=1}^{l} \gamma_i \xi_i 
\]

(5)

Translates into the solution to its dual problem:

\[
\left\{ \begin{array}{l}
\max \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} y_i y_j \alpha_i \alpha_j (x_i^T, x_j) \\
\text{s.t} \quad 0 \leq \alpha_i \leq C, \quad i = 1, 2, \ldots, l \\
\sum_{i=1}^{l} \alpha_i y_i = 0
\end{array} \right.
\]

(6)

Find a set of \( \alpha \) that satisfies the above conditions and solve for \( w \) and \( b \).

\[
w = \sum_{i=1}^{l} \alpha_i y_i x_i
\]

(7)

\[
b = \max_{k, j} \frac{w^T x_i + \min_{k, j} w^T x_j}{2}
\]

(8)

When the training set data is linearly inseparable, it can be mapped into a high-dimensional space by a nonlinear mapping, making it linearly separable in the high-dimensional space. The vector \( x \) is replaced by \( \psi(x) \) and the inner product \( (x_i^T, x_j) \) is replaced by \( K(x_i, x_j) = \psi(x_i)^T \psi(x_j) \). \( K(x_i, x_j) \) is called the kernel function, it is to calculate the vector inner product value in the transformed high-dimensional space after inputting the low-dimensional space vector. In SVM, there are the following four commonly used kernel functions:

\[
\begin{align*}
\text{linear}: & \quad K(x_i, x_j) = x_i^T \cdot x_j \\
\text{polynomial}: & \quad K(x_i, x_j) = (a x_i^T x_j + b)^k, \quad a > 0 \\
\text{RBF}: & \quad K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \quad \gamma > 0 \\
\text{sigmoid}: & \quad K(x_i, x_j) = \tanh(ax_i^T x_j + b)
\end{align*}
\]

(9)
4. Assessment model

The assessment model is shown in Figure 2 and is divided into 3 steps.

Step 1: Collect historical data of power systems under different conditions, and build 36-dimensional original features as training data sets based on the literature [2, 3].

Step 2: Using the SAE network structure, unsupervised learning from the original features and obtaining robust high-order representation features.

Step 3: Add the obtained characterization features as input to the SVM for training and testing.

In this paper, the confusion model is used to comprehensively evaluate the assessment model. As shown in Table 1, Class=1 and Class=-1 are respectively expressed as the situation of stability and instability. TP and TN are the correct evaluation results. FN means the prediction is unstable but the system is actually stable. FP means the system is transiently unstable but the prediction is transiently stable.

### Table 1. Confusion Matrix.

| Class in reality | Class=1 | Class=-1 |
|------------------|---------|----------|
| Class=1          | TP      | FN       |
| Class=-1         | FP      | TN       |

The accuracy rate PACC, the false dismissal rate PFD, and the false alarm rate PFA were used as the evaluation indexes of the classification performance.

\[
P_{\text{ACC}} = \frac{TP + TN}{TP + FN + FP + TN}
\]  
(10)

\[
P_{\text{FD}} = \frac{FP}{TP + FN + FP + TN}
\]  
(11)
When applied to the online stability assessment of the power system, the online data can be obtained from the PMU to construct the original feature to input to the already trained assessment model and determine whether the power system is transiently stable.

5. Simulation analysis
The example uses the New England 39-bus system, the generators use fourth-order model, and all loads use a constant-impedance model. Consider 80%, 85%,... , 120% for 9 load levels, correspondingly change the generator output. The fault is set to three-phase short circuit, the fault is applied at 0.1s and cleared at 0.35s. The simulation time is set to 4s. The instability criterion is set as the maximum power angle difference of the generator is more than 180°. The simulation software is PSAT based on Matlab platform. A total of 1560 samples are obtained by simulation. Among them, 1170 samples are used as training data sets, and the remaining 390 samples are used for testing.

The appropriate parameters of the model are obtained through debugging. The test results of the proposed method and the other two methods are shown in Table 2.

| Model          | PA, %  | PD, %  | PA, %  |
|----------------|--------|--------|--------|
| SVM            | 97.44  | 1.28   | 1.28   |
| PCA+SVM        | 97.18  | 1.79   | 1.03   |
| SAE+SVM        | 97.94  | 1.03   | 1.03   |

It can be seen from Table 2 that the SAE feature extraction model improves the evaluation accuracy compared with other models, while the false alarm rate and false dismissal rate are slightly reduced, and the evaluation performance is better. SAE has superior feature extraction capabilities compared to PCA. This is because the dimensionality reduction function of PCA is based on linear transformation, and the abstraction ability of the underlying measurement data is limited. SAE constructs a deep structure with the goal of minimizing reconstruction error, which can retain the underlying input information to the maximum extent and extract features layer by layer.

6. Conclusion
This paper presents a new transient stability assessment model. The original input features are constructed by the PMU data collected in the wide-area measurement system, and then the multi-layer abstract learning is used to extract key features using SAE, and the SVM is used as a classifier to determine whether the transient is stable. The simulation was performed in the New England 39-bus system. The simulation results show that the proposed evaluation model shows satisfactory performance and is suitable for real-time online identification of power system transient stability.

In this paper, we only initially explored the application of the deep learning model in the power system. Later, we will explore more assessment models to apply to the complex and variable actual power grid.

References
[1] Shin H C, Orton M R, Collins D J, et al. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35, (8), pp. 1930-1943.
[2] Shengyong Y, Xiaoru W, Zhigang L, et al. Dual-stage feature selection for transient stability assessment based on support vector machine, Proc. CSEE, 2010, 30, (31), pp. 28–34.
[3] Fei T. Research on feature extraction and assessment method of power system transient stability based on WAMS. 2013.

\[
P_{\text{fa}} = \frac{FN}{TP + FN + FP + TN} \quad (12)
\]
[4] Tang Y, Yang J. Dynamic event monitoring using unsupervised feature learning towards smart grid big data. Int. Joint Conf. Neural Networks, 2017.

[5] Savulescu S C. Real-time stability in power systems: techniques for early detection of the risk of blackout. New York, 2006.

[6] Sun K, Likhate S, Vittal V, et al. An online dynamic security assessment scheme using phasor measurements and decision trees. IEEE Transactions on Power Systems, 2007, 22, (4), pp. 1935-1943.