applied sciences

MDPI

Article

An Improved Power System Transient Stability Prediction Model Based on mRMR Feature Selection and WTA Ensemble Learning

Jun Liu [1,*], Huiwen Sun [1], Yitong Li [1], Wanliang Fang [1,*] and Shuanbao Niu [2]

[1] Shaanxi Key Laboratory of Smart Grid, School of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710049, China; shw662064@stu.xjtu.edu.cn (H.S.); niu2282@stu.xjtu.edu.cn (Y.L.)
[2] Northwest Subsection of State Grid Corporation of China, Xi'an 710048, China; lasa0918@163.com
* Correspondence: eeliujun@mail.xjtu.edu.cn (J.L.); eewlfang@mail.xjtu.edu.cn (W.F.); Tel.: +86-29-8266-8782 (W.F.)

Received: 17 February 2020; Accepted: 21 March 2020; Published: 26 March 2020

Abstract: Fast online transient stability assessment (TSA) is very important to maintain the stable operation of power systems. However, the existing transient stability assessment methods suffer the drawbacks of unsatisfactory prediction accuracy, difficult applicability, or a heavy computational burden. In light of this, an improved high accuracy power system transient stability prediction model is proposed, based on min-redundancy and max-relevance (mRMR) feature selection and winner take all (WTA) ensemble learning. Firstly, the contributions of four different series of raw sampled data from all of the three-time stages, namely the pre-fault, during-fault and post-fault, to transient stability are compared. The new feature of generator electromagnetic power is introduced and compared with three conventional types of input features, through a support vector machine (SVM) classifier. Furthermore, the two types of most contributive input features are obtained by the mRMR feature selection method. Finally, the prediction results of the electromagnetic power of generators and the voltage amplitude of buses are combined using the WTA ensemble learning method, and an improved transient stability prediction model with higher accuracy for unstable samples is obtained, whose overall prediction accuracy would not decrease either. The real-time data collected by wide area monitoring systems (WAMS) can be fed into this model for fast online transient stability prediction; the results can also provide a basis for the future emergency control decision-making of power systems.

Keywords: transient stability prediction; electromagnetic power; SVM; mRMR feature selection; WTA ensemble learning

1. Introduction

Social developments and economic growths have been calling for higher requirements for secure and reliable power supplies. With the continuous and rapid growth of electricity demand, the power systems are developing towards large-scale [1,2], high voltage, regional grid interconnection, hybrid AC/DC, long-distance large-capacity transmission and high renewable energy penetration [3–5], etc., The power system topologies and operation characteristics are becoming increasingly complex and changeable, which has brought severe challenges to the secure and stable operation of power systems. The challenges of non-linearity and the rapid development of the electromechanical transient process make it difficult to predict the system transient stability quickly and accurately after the fault has been cleared [6]. In recent years, many large-scale power system blackouts have occurred worldwide, which makes fast-online transient stability assessments and emergency control more urgent [7–10].

Traditional transient stability assessment (TSA) methods include a time-domain method based on (1) electromechanical transient simulations [11–13], and (2) direct methods based on Lyapunov's stability theory [14,15], transient energy function (TEF) related approaches [2,16], equal area criterion (EAC) related approaches [9,17], and so on. Time-domain simulation with detailed dynamic component models is more accurate and can be used as a standard for testing other TSA methods. However, simulation-based methods are often used offline, because of the large number of calculations required when considering large-scale power systems and large contingency sets. Direct methods can reflect the transient characteristics of power systems, by constructing an appropriate Lyapunov energy function, which is a sufficient but unnecessary condition for TSA, with a strict mathematical foundation. However, it is usually rather difficult to formulate an energy function that satisfies the complex operation conditions of a nonlinear power system. Therefore, the two traditional TSA methods cannot meet the requirements of fast and accurate online transient stability assessment.

Since the late 1980s, with the rise of data mining techniques, Louis Wehenkel et al. began to apply machine learning (ML) algorithms such as the decision tree (DT) and artificial neural network (ANN) to power system transient stability predictions [18]. The key issue to transient stability prediction is to establish the mapping relationship between system features X and transient stability y that is, to find the nonlinear function of $y = f(X)$. In order to learn this mapping relationship, a large number of samples need to be obtained through offline time-domain simulations; and this nonlinear function can then be fitted using the ML algorithm. Due to the rapid development of machine learning and its related mathematical theories, ML-based TSA research has been widely carried out worldwide; (1) ANN related models, including adaptive ANN application for dynamic security assessment [19], convolutional neural networks [20,21], deep imbalanced learning framework [22], and deep belief network and model interpretation method [23] for transient stability assessment; (2) DT related models to predict system vulnerability [24], and transient instability [25]; (3) support vector machines (SVM) related models, such as an improved SVM, were proposed for real-time TSA in power systems in Reference [26]. Rotor speed, rotor angle of generators and voltage amplitude of buses after fault were extracted as the input features to train and test the SVM model in Reference [27]. A two-stage feature selection method based on SVM was presented In Reference [28]. Reference [29] introduced the information fusion technology into TSA, in order to obtain feature sets with reduced dimensions. Additionally, a variant of SVM, like core vector machine (CVM), was also proposed to solve the TSA problem based on phasor measurement units (PMUs) big data [30]; (4) other ML models, such as the extreme learning machine (ELM) algorithm [31], the least absolute shrinkage and selection operator (LASSO) [32], etc., have also been applied to the transient stability prediction of power systems.

Among all ML algorithms, SVM is one of the most commonly used methods in transient stability prediction, because the TSA problem can be seen as a simple binary classification problem. However, most of the existing research that applies SVM to the transient stability prediction only extracted the trajectory variables after fault as input features, which did not fully consider the dynamic characteristics of the power system at different time stages. However, some ensemble learning models have been established to improve prediction performance, for example, the data segmentation-based ensemble classification (DSEC) method was proposed for transient stability status prediction with imbalanced data [6], the transient stability prediction by a hybrid intelligent system [33], and so on. The prediction accuracies of existing models using those post-fault input features are usually not high enough, which might not be suitable for subsequent online emergency control.

In view of the above concerns, the contribution of this paper includes the following four aspects. Firstly, the influence of trajectory variables in multiple time stages on transient stability is analyzed, and the sampled data covering all three-time stages, namely the pre-fault, during-fault and post-fault, are proposed to be used as the SVM input features. Secondly, according to the rotor motion equation (also known as the "swing equation"), the new feature of generators' electromagnetic power is introduced into the transient stability prediction. Case analysis shows that generators' electromagnetic power has higher classification accuracy than other feature types. Then, further feature selection is

carried out based on min-redundancy and max-relevance (mRMR), and the optimal feature set for predicting system stability is obtained. Finally, the prediction results of SVM classifiers with different feature sets are combined by the winner take all (WTA) method, to establish a high-precision and conservative transient stability prediction model.

2. Materials and Methods

Power system stability can be classified into rotor angle stability, voltage stability, and frequency stability. Rotor angle stability can be further divided into small-disturbance stability and transient stability. The definitions and classifications can also be found in Reference [30]. Since transient stability assessment is one of the most important issues to guarantee the secure and stable operation of power systems, it belongs to short-term rotor angle stability ramification [34]. Therefore, only the rotor angle stability in the power system electromechanical transient process is studied, and the voltage stability, frequency stability and medium-and-long term stability are not considered. The improved power system transient stability prediction model can be realized through the following four steps: data preparation, multi-input feature analysis, mRMR feature selection and WTA ensemble learning modeling.

2.1. Data Preparation and Basic SVM Prediction Model

2.1.1. Three-time Stages Related to Transient Stability

The electromechanical transient stability of the power system is related to three-time stages, pre-fault (steady state), during-fault and post-fault, as shown in table Figure 1. The pre-fault stage reflects the initial operation state of the system, the during-fault stage reflects the severity of the fault disturbance, and the post-fault stage reflects the dynamic performance of the system after the fault is cleared.

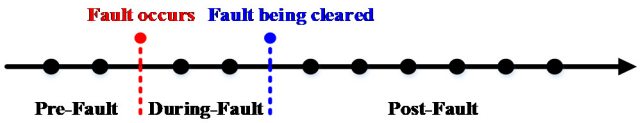


Figure 1. Three-time stages related to transient stability.

Since the power system is a complex non-linear system, its transient stability is not only related to the post-fault information, but also influenced by the steady-state operation point of the power system and the severity of disturbance during the fault. In order to analyze the different contributions of the variables during these three time stages to the system transient stability, the three time stages are grouped into 4 series of sampling data, as shown in Figure 2.

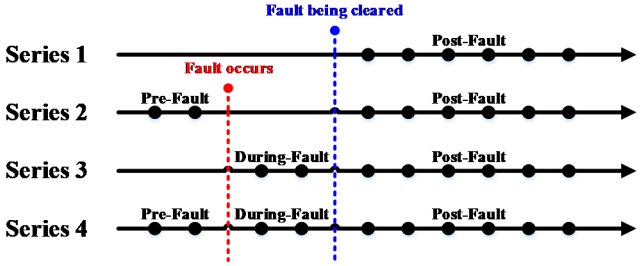


Figure 2. Different sampling series.

Comparisons on these 4 series of sampled data are tested on the IEEE 39-bus system, which is shown in Figure 3, using the dual-axis generator model, IEEE DC Exciter Type 1 exciter model, and constant-impedance load model. The simulation software is MATLAB toolbox PST3.0 [35].

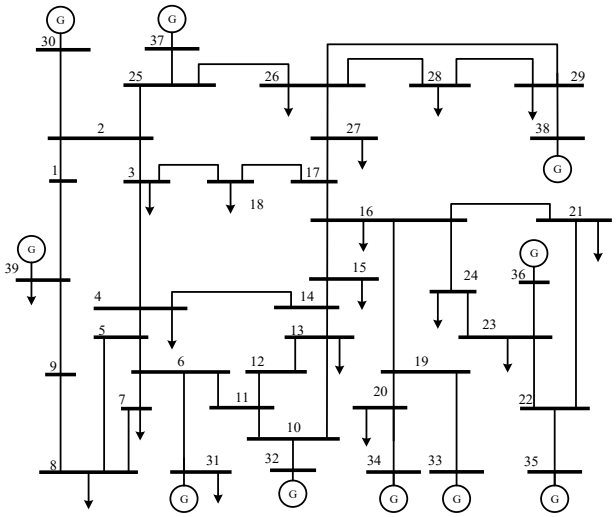


Figure 3. Topology of the IEEE 39-bus system.

The transient instability criterion is set as, if the maximum rotor angle difference between any pair of generators exceeds 180° at the end of the transient simulation (such as 5s), the case will be recognized as unstable [36]. Figure 4 shows the bus voltage amplitude curves of a typical transiently stable sample (in blue) and an unstable sample (in red), with the four series of sampled data. Here, the fault occurred at 0.10 s and was cleared at 0.20 s in all cases. For the stable sample, the load is 0.9 times the base load level, the faulted line is 1–39, and the fault location is 10% from bus-39 side. For the unstable sample, the load is 1.1 times the base load level, the faulted line is 21–22, and fault location is 40% from the bus-21 side.

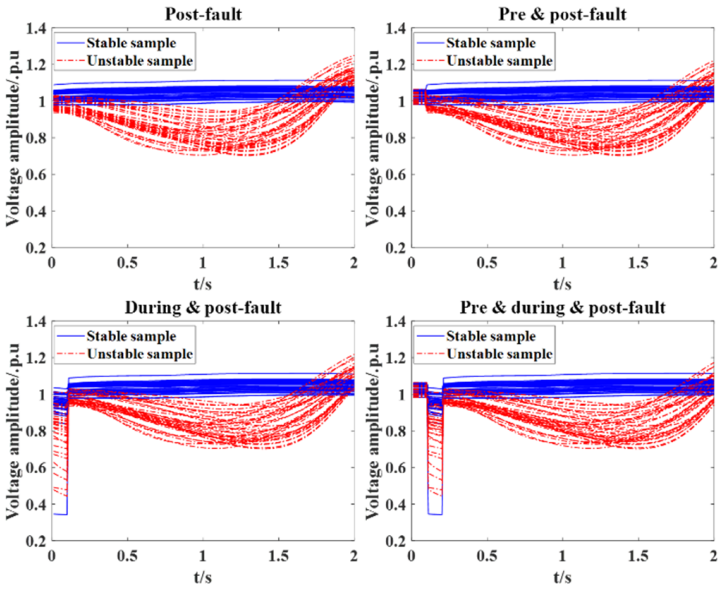


Figure 4. Voltage curves of all buses for the four-sampling series.

It can be seen from Figure 4 that the first three series cannot fully reflect the exact dynamic behavior of the power system, due to the loss of information. In order to better characterize power system transient performance, it is necessary to utilize the sampling data covering all three time stages, that is, the stages of pre-fault, during-fault and post-fault, as shown in Series 4 of Figure 2. It should also be noted that in actual power systems, the typical sampling frequency for the fundamental phasor of the trajectory variables is 10 ms for the PMUs in the wide area measurement system (WAMS).

2.1.2. SVM Prediction Model

It can be seen that the input features of Figure 4 are not linearly separable, because some data points of the stable and unstable samples of the voltage curves intersect with each other. SVM is able to map a linearly inseparable data in low-dimensional space to a linearly-separable high-dimensional space through kernel functions. Figure 5 shows the mapping process visually.

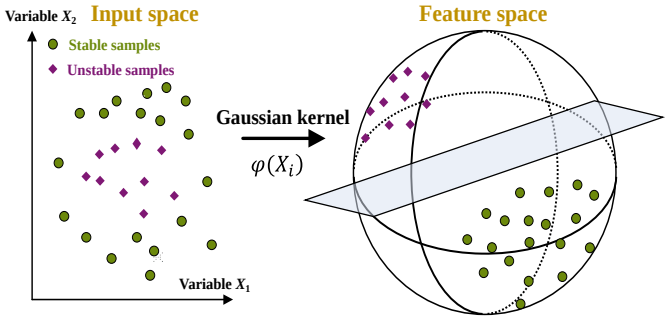


Figure 5. Principal diagram of SVM kernel function.

The basic principle of SVM is shown in Equation (1).

$$\left\{ \begin{array}{l} \min\limits_{w,b,\zeta}\left(\frac{1}{2}\right)w^T w + C\sum\limits_{i=1}^{N}\zeta_i \\ s.t.\left(w^T\varphi(X_i)+b\right)yi \geq 1-\zeta_i \\ \zeta_i \geq 0, i = 1,\ldots N \\ f(X) = \sum\limits_{i=1}^{N}\alpha_i y_i \varphi(X_i)^T\varphi(x)+b \end{array} \right\}, \tag{1}$$

where ω is the weight vector of the hyperplane; b is the threshold value; ζ is the relaxation variable; C is the penalty factor for the relaxation variable; N is the number of training samples; $\varphi(.)$ is the mapping function from low-dimensional space to high-dimensional space, the kernel function is chosen as radial basis function (RBF); $X_i(i = 1,\ldots,N)$ are the support vectors; y_i is the output of i_{th} sample; $f(.)$ is the fitted model of SVM, where α_i is the coefficient of i_{th} sample.

2.2. Analysis of Multiple Input Features

In the electromechanical transient stability analysis, the transient stability criterion is determined by the maximum rotor angle difference between each pair of generators. The rotor angle dynamic behavior can be influenced by many factors, and the detailed feature extraction for transient stability assessment is analyzed below.

2.2.1. Rotor Motion Equation

According to the rotor motion Equation (2), the rotor angle of generators is directly affected by the rotor speed, electromagnetic power and mechanical power of generators.

$$\left\{ \begin{array}{l} \frac{d\delta}{dt} = \omega - 1 \\ T_J\frac{d\omega}{dt} = \frac{P_m - P_e}{\omega} - D\omega \end{array} \right. , \tag{2}$$

where δ is the rotor angle; ω is the rotor speed; P_m is the mechanical power; P_e is the electromagnetic power; T_j is the inertia time constant; and D is the damping coefficient of the generator.

The electromagnetic power can be calculated by Equation (3), during the time domain simulation.

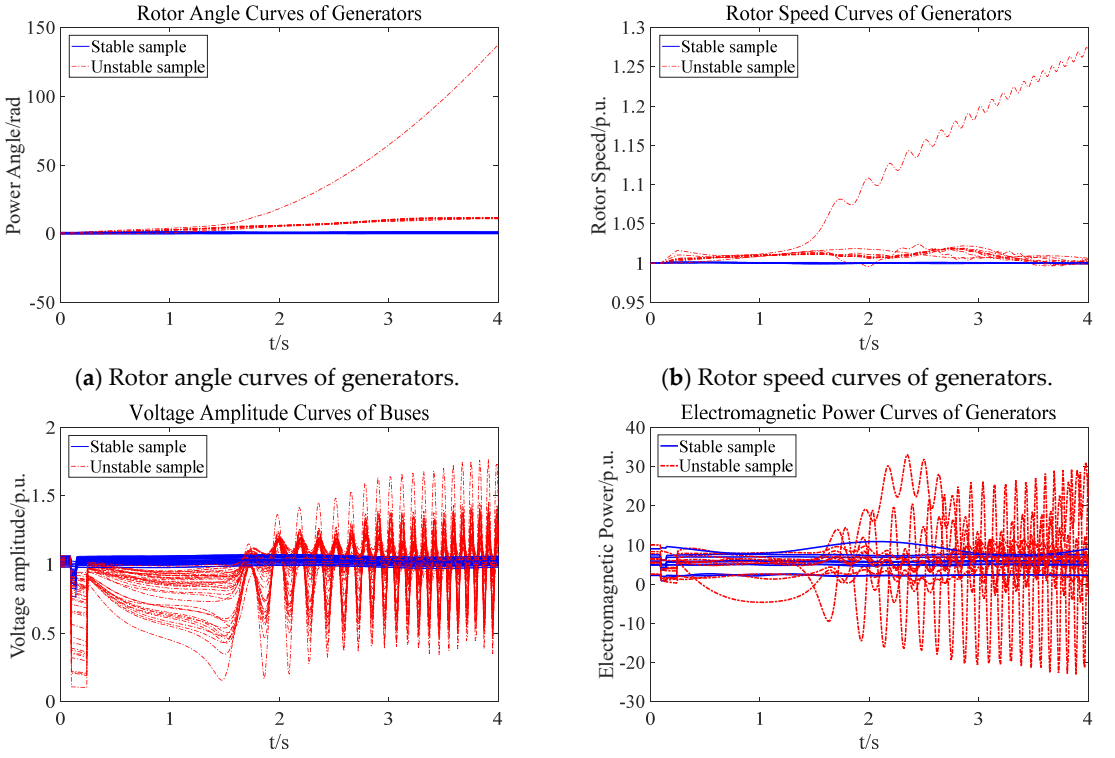
$$Pe = V_d I_d + V_q I_q, \tag{3}$$

where V_d and V_q are the d-axis voltage and q-axis voltage, respectively; I_d and I_q are the d-axis current and q-axis current, respectively.

Due to the rapid development of the transient process, it is assumed that the mechanical power P_m does not change during the short process. The damping coefficient D is also neglected in this study. When a certain fault occurs, the electromagnetic power P_e of generators will change rapidly with the changes of voltage and current variables; the imbalance power between P_m and P_e will then cause the rotor speed ω to change. Rotor speed ω changes will further affect the rotor angle δ, that is, the transient stability of power system can be partially traced back to P_e. Therefore, the electromagnetic power can also be considered as an important factor affecting the transient stability besides the existing input features [27], namely, the rotor angle δ of generators, rotor speed ω of generators and voltage amplitude V of buses.

2.2.2. Separability of Electromagnetic Power and 3 Traditional Features

Based on the analysis above, four types of input features [δ,ω,V,P_e] can be used to train the SVM transient stability prediction model. Figure 6 shows the simulation curves of the 4 types of features, extracted from typical stable and unstable samples of the IEEE 39-bus system.



(**a**) Rotor angle curves of generators.

(**b**) Rotor speed curves of generators.

(**c**) Voltage amplitude curves of buses.

(**d**) Electromagnetic power curves of generators.

Figure 6. Simulation curves of the 4 types of features for stable and unstable samples.

It can be seen from Figure 6d that the transiently stable and unstable samples of the P_e curves are significantly different, similar to the δ, ω and V features in Figure 6a–c. Therefore, it is feasible to use P_e as a novel input feature.

2.3. Feature Selection Based on mRMR Technique

In large-scale power systems, too many input features of the SVM model will cause a heavy computational burden, which limits the online applicability of the machine learning model. After briefly comparing the 4 types of input features, further feature selection based on mRMR will be conducted in this section.

It is obvious that the transient stability prediction time increases with the increase of the feature size, as shown in Figure 7 (using the IEEE 39-bus system). Therefore, when combining multiple features into the SVM classifier, it is necessary to reduce the number of input features.

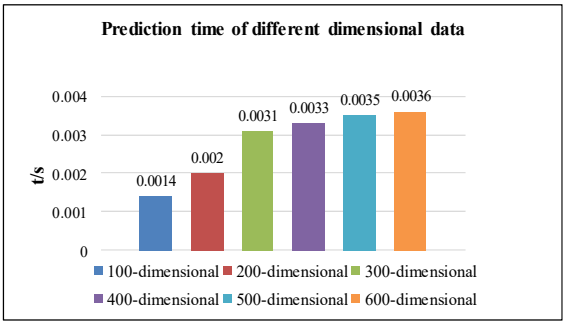


Figure 7. Prediction time of different dimensional data.

In the field of information theory, mutual information is always used to measure the degree of correlation between discrete random variables, as shown in Equation (4).

$$\begin{aligned} I(X;Y) &= \sum_{x\in x_1}\sum_{y\in x_2} p(x,y)\log\frac{p(x,y)}{p(x)p(y)}, \\ I(X,Y) &= I(Y;X) \end{aligned} \tag{4}$$

where $p(x,y)$ is the joint probability density function of X and Y; x_1, x_2, $p(x)$ and $p(y)$ are the value spaces and marginal probability density functions of X and Y, respectively. Unless otherwise specified, the subscript of the logarithm is 2.

Mutual information calculation requires the variables are discrete, so it is necessary to discretize the trajectory variables with continuous values first. In order to make the SVM learning model have better fitting performance, the data are normalized to the range of [0,1] in the meantime.

As mentioned in Section 2.1.1, trajectory variables are used as input features; each group of input features contains 10-dimensional discrete data. Although each dimension of the data contains certain information, if the mutual information analysis is performed directly on each dimension of the data using Equation (4), the original physical meaning of trajectory variables will be destroyed. Therefore, each series of trajectory variables, namely, the 10-dimensional trajectory data, are regarded as a single group of input features for mutual information calculation, due to the temporal correlation among them. Taking the voltage amplitude of buses as an example, for the bus i and bus j, the sampled 10-dimensional voltage amplitude data are regarded as two vectors, as shown in Equation (5).

$$\begin{aligned} busi: V_i &= \left[{V_i}^{(1)},\ldots,{V_i}^{(10)}\right], \\ busj: V_j &= \left[{V_j}^{(1)},\ldots,{V_j}^{(10)}\right] \end{aligned} \tag{5}$$

Thus, a joint probability distribution of the two groups of input features is proposed for mutual information calculation of the time-correlated trajectory data, as shown in Equation (6).

$$I(Vi;Vj) = \sum_{m=1}^{10}\sum_{n=1}^{10} p\left({V_i}^{(m)},{V_j}^{(n)}\right)\log\frac{p\left({V_i}^{(m)},{V_j}^{(n)}\right)}{p\left({V_i}^{(m)},{V_j}^{(n)}\right)}, \tag{6}$$

Based on the definition of mutual information, the feature selection method mRMR [37] is utilized to obtain an optimal feature subset, which has minimum redundancy among the interior features and maximum relevance with the stability result.

For a dataset $D = \{x_1, \ldots, x_N | y\}$ with N groups of input features and the stability label vector y. Assuming that S is a subset of D, the redundancy of the subset can be calculated as Equation (7).

$$V_s = \frac{1}{|S|^2} \sum_{i,j \in S} I\left(x_i; x_j\right), \tag{7}$$

where $|S|$ is the number of feature groups contained in the subset S.

The correlation between subset S and target vector y is calculated as Equation (8).

$$W_s = \frac{1}{|S|} \sum_{i \in S} I(x_i; y), \tag{8}$$

Then, taking Equation (9) as the optimization objective to find the optimal subset of input features, with less redundancy V_S and stronger relevance W_S.

$$\max_S (W_S / V_S), \tag{9}$$

The computational burden of obtaining the optimal feature subset is huge. Thus, mRMR technology uses the incremental search algorithm to sort all feature groups and then select the optimal feature subset. The detailed process is as follows.

1. Define the set of selected feature groups as S.
2. Calculate the correlation between each group of input features x_1 and the target y, and then select the group of input features $x^{(1)}$ that is most relevant to the target according to Equation (10). The selected group of input features $x^{(1)}$ is added to the set S as the first input feature group.

$$\max_{X^{(1)} \in D} I\left(x^{(1)}; y\right), \tag{10}$$

3. Select the next group of input features $x^{(j)}$ according to Equation (11), using the previously recorded features $x^{(i)}$ in S.

$$\max_{x^{(j)} \in D - S} \left[I\left(x^{(j)}; y\right) / \left(\frac{1}{|S|} \sum_{x^{(i)} \in S} I\left(x^{(j)}; x^{(i)}\right) \right) \right], \tag{11}$$

4. Add the selected feature group $x^{(j)}$ in Step 3 to the set S, and then repeat Step 3 until all input features are sorted.

The final sorting result indicates that if a subset of $N_1 (N_1 \leq N)$ feature groups are selected as the input of the learning machine, the first N_1 feature groups in the set S will be the optimal subset, which shows a stronger correlation with the target and less interior redundancy.

Since the bus V has a total of 39 groups of features, and the generator P_e had only 10 groups of features, the bus V seems to make a greater contribution to the transient stability prediction in mRMR feature selection. On the other hand, the transient stability prediction results in Section 3.2 show that the classification accuracy of the generator P_e is higher than that of the bus V. However, both types of features have better prediction results than the electromechanical variables of δ and ω. Section 2.4 will aim to combine these two types of superior features V and P_e, and jointly form an improved transient stability prediction model.

2.4. High Accuracy Prediction Model Based on WTA Ensemble Learning

As shown in Section 3.2, although the prediction model of the new feature of generator P_e reaches an accuracy of about 98.77%, it is necessary to establish a more accurate transient stability prediction

model, especially to accurately identify the unstable situations, in order to avoid losing synchronization, cascading failures, or even large-scale blackout.

2.4.1. Combined Features of Voltage Amplitude and Electromagnetic Power

In order to meet the fast and accurate prediction requirement online, the two types of superior features with higher accuracy are selected according to the mRMR ranking result in Section 2.3, namely the generator electromagnetic power P_e and bus voltage magnitude V.

2.4.2. WTA Ensemble Learning Model

As analyzed above, the overall prediction accuracies of P_e and V are relatively high, but the prediction accuracies for unstable samples are still very low. In order to improve the conservativeness of the prediction model, the SVM learning machines based on P_e and V are taken as two sub-classifiers. Then the outputs of the two classifiers are combined by the winner take all (WTA) ensemble learning method. When the prediction result of any sub-learning machine is unstable, the WTA module will determine that the transient process is unstable; otherwise it will be accepted as stable. The input features of sub-learning machine 1 (M1) is V, and the input features of sub-learning machine 2 (M2) is P_e. Stable samples have a label of 1 and unstable samples have a label of 0. The principle of WTA ensemble learning model is expressed as Equation (12).

$$\begin{cases} y_1 = f_1(V) \\ y_2 = f_2(P_e) \end{cases} \\ y = \min\{y_1, y_2\} \tag{12}$$

The WTA ensemble learning process can also be shown in Figure 8.

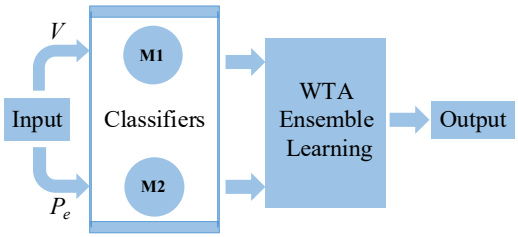


Figure 8. WTA ensemble learning model.

As previously mentioned, the V and P_e curves reflect the different dynamic characteristics of power systems. In other words, the transient stability prediction results of M1 and M2 can be relatively independent. Assuming that the error rates of M1 and M2 are ε_1 and ε_2 respectively, then the error rate of the WTA model for unstable samples is shown in Equation (13).

$$\varepsilon_{WTA} = \varepsilon_1 \times \varepsilon_2 \ll \min\{\varepsilon_1, \varepsilon_2\}, \tag{13}$$

For stable samples, the prediction accuracy P_{st} of the WTA model is the product of the accuracies ${P_{st}}^{M1}$ and ${P_{st}}^{M2}$. For unstable samples, the prediction accuracy P_{um} of the WTA model is 1 minus the error rate product of the unstable predictions $\left(1 - {P_{um}}^{M1}\right)$ and $1 - {P_{um}}^{M2}$. Therefore, the overall prediction accuracy P_{total} of the WTA model can be calculated by 1, minus the proportion of the number of samples with incorrect prediction results to the total number of samples $(N_{st} + N_{um})$, as shown in Equation (14).

$$\begin{cases} P_{st} = {P_{st}}^{M1} \times {P_{st}}^{M2} \\ P_{um} = 1 - \left(1 - {P_{um}}^{M1}\right) \times \left(1 - {P_{um}}^{M2}\right) \\ P_{total} = 1 - \frac{N_{st} \times (1 - P_{st}) + N_{un} \times (1 - P_{um})}{N_{st} + N_{un}} \end{cases}, \tag{14}$$

3. Results

In order to verify the methods in Section 2, the sample series analysis, input feature extraction, mRMR feature selection, and transient stability prediction results based on WTA ensemble learning are described respectively as follows.

3.1. Sample Generation and Data Series Analysis with Traditional Three Features

The process of generating the simulated data samples is as follows. The load level is randomly set to 0.9, 1.0 or 1.1 times of the base load (the generator outputs are adjusted proportionally). Three-phase short-circuit faults are applied on the selected 33 transmission lines (excluding transformers and islands), at 10% to 80% positions, with an interval of 10%. The circuit breakers might trip the faulted line at 0.05 s, 0.1 s, 0.15 s or 0.20 s after the fault occurrence. A total of 3168 samples are obtained, of which two-thirds are randomly selected as training samples, and the rest are test samples.

The sampling time step is 0.01 s. The sampled data of all three-time stages include two consecutive points before the fault, two consecutive points immediately after the fault occurrence, and 6 consecutive points after the fault is cleared. The system variables original selected as input features are rotor angles of 10 generators, rotor speeds of 10 generators and voltage amplitudes of 39 buses, all in units of per unit (p.u.). To verify the effectiveness of the corresponding four sampling series in Section 2.1.1, the basic prediction results of these three types of input features are compared in Figure 2.

In this paper, the SVM model used is libsvm2.0 [38], and the optimal penalty coefficients C and RBF kernel parameters of the SVM classifiers are obtained by grid search [36] and 5-fold cross-validation using the training samples. Then, the entire training samples and the optimal SVM parameters are used to retrain the SVM classifiers. Finally, the prediction accuracy of SVM classifiers is obtained by testing the test samples. The results for these three types of input features from all four sampling series are shown in Figure 9.

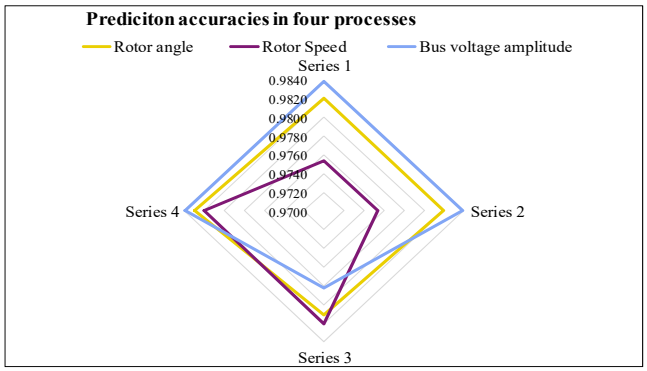


Figure 9. Prediction accuracies of the 4-sample series.

It can be seen from Figure 9 that Series 4 has the highest prediction accuracy among all 4 series, which is greater than 98.00% for all three types of input features; while the accuracies are lower than 97.35%, 97.35% and 97.60% for Series 1, 2 and 3, respectively. This means that the feature set that contains all of the pre-fault, during-fault and post-fault stages can better characterize the transient behavior of the system. Therefore, in the following subsections, the input features will include the 10 data points (sampled like Series 4) of each trajectory variable from all three-time stages.

3.2. Prediction Results of Four Input Features, Including the Proposed Electromagnetic Power Feature

As described in Section 3.1, the optimal penalty coefficients and RBF kernel parameters of the SVM classifiers can be acquired through the training samples with features δ, ω, V and P_e, respectively. Theses 4 independent SVM classifiers can then be established to estimate the prediction accuracy of different types of input features on the test samples. The results are shown in Table 1.

Table 1. Prediction performance of different types of input features.

Electrical Quantity	Train Samples	Test Samples	Train Accuracy/%	Test Accuracy/%
Generator current I	2112	1056	99.29	97.73
Rotor angle δ			99.29	98.30
Rotor speed ω			99.15	98.20
Bus voltage amplitude V			99.95	98.39
Electromagnetic power P_e			99.62	98.77

It can be seen in Table 1 that, among the 4 types of input features, the transient stability prediction accuracy of the feature P_e is the highest, reaching 98.77%, followed by the feature V, etc., The features of δ and ω have relatively lower accuracies because they may not change much within the 0.06 s-period after the fault clears, namely only 6 post-fault sampling points.

The prediction results of SVM are further compared with Back Propagation Neural Networks (BP-NN) and Random Forest (RF) in Table 2; the BP-NN algorithm is from [39], and RF uses the built-in function of MATLAB. It can be seen in Table 2 that the accuracy of SVM is higher than that of BP-NN and RF. Therefore, the SVM sub-classifiers are mainly used in this study.

Table 2. Confusion Matrix for classification results of electromagnetic power.

Model	Optimal Parameters	Electrical Quantity	Test Accuracy/%
SVM	5-fold: C = 2048, g = 0.25	Bus voltage amplitude V	98.39
	5-fold: C = 512, g = 0.002	Electromagnetic power P_e	98.77
BP-NN	The iteration epochs: 100 The batch size: 352	Bus voltage amplitude V	97.35
		Electromagnetic power P_e	98.01
RF	500 trees, 20 random variables	Bus voltage amplitude V	97.82
	500 trees, 10 random variables	Electromagnetic power P_e	96.97

Tables 3 and 4 further give the confusion matrix of the bus voltage amplitude and the electromagnetic power of the generators, which shows the prediction results of the SVM classifier in more detail. In the tables, the recall rates denote the proportions of transiently stable samples and unstable samples that were accurately predicted, separately.

Table 3. Confusion Matrix for classification results of bus voltage amplitude.

Test Sets Labels	Predict Results		Recall Rate/%
	Stable	Unstable	
Stable	945	7	99.26
Unstable	10	94	90.38

Table 4. Confusion Matrix for classification results of electromagnetic power.

Test Sets Labels	Predict Results		Recall Rate/%
	Stable	Unstable	
Stable	947	5	99.47
Unstable	8	96	92.31

It can be seen in Tables 3 and 4 that the recall rates of stable samples are relatively high, always above 99%, but the recall rates of the unstable samples are still not high enough. Therefore, the rest of this paper will focus on improving the prediction performance of unstable samples, that is, to improve the conservativeness of the model.

3.3. Optimal Input Features Selection by mRMR

The variables δ, ω and P_e of 10 generators and the V of 39 buses constitute of a 690-dimensional feature set. The 69 groups of trajectory variables are arranged in the following order: δ numbered 1–10, ω numbered 11–20, V numbered 21–59, and P_e numbered 60–69. The basic result sorted by mRMR is shown in Figure 10.

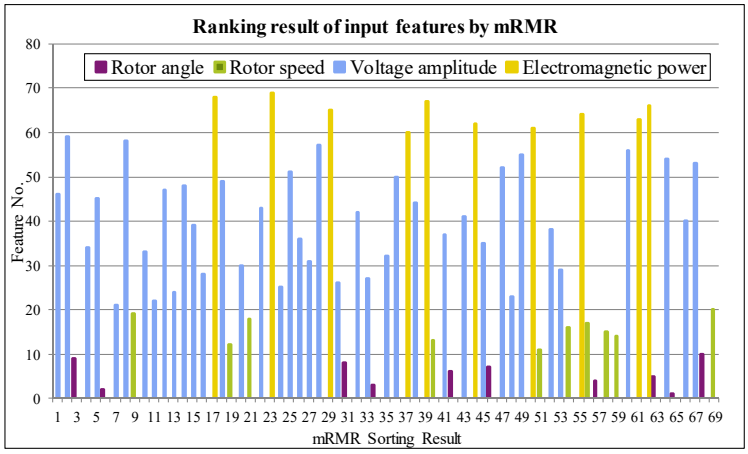


Figure 10. Sorting result of input features by mRMR.

It can be seen from Figure 10 that 7 out of the top 10 groups of features most relevant to transient stability are bus voltage variables V (in blue), which can be seen as a good feature type to assess the transient stability of power systems.

However, bus V has 39 groups of input features, which is far more than the number of generator variables δ, ω and P_e. Therefore, in order to analyze the priorities of the other three types of generator-related features to transient stability, it is necessary to temporarily remove bus V data from the ranking results. Then the numerical order of the generator-related features δ, ω and P_e are respectively added, and the sums are shown in Figure 11.

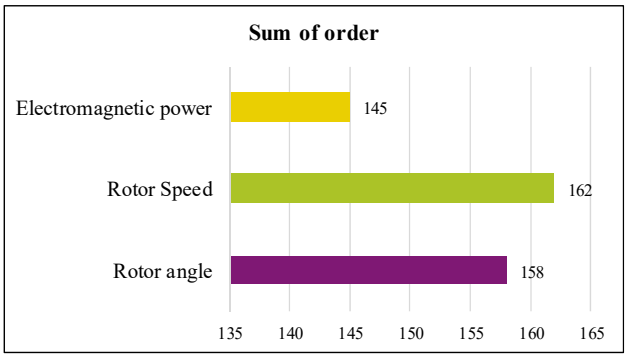


Figure 11. Sum of the order in three types of features.

As can be seen from Figure 11, the order sum of the P_e (in yellow) group is far less than those of the variables δ and ω. In other words, the new feature of P_e can be acknowledged as more relevant to the system transient stability than δ and ω, which is also consistent with the accuracy comparisons of transient stability prediction in the previous Section 3.2.

3.4. High Accuracy Prediction Results Based on WTA Ensemble Learning

3.4.1. Simple Combined Features of Voltage Amplitude and Electromagnetic Power

The changes in electromagnetic power P_e will cause changes in the mechanical performance of generators, while the trajectories of bus voltage V can reflect the dynamic voltage recovery of the power system. An intuitive combination method is to integrate these two types of features into a single SVM learning machine, using a total of 10 + 39 = 49 trajectory variables, each variable containing 10-dimensional data from the pre-fault to post-fault stages. The prediction results are shown in Table 5 and Figure 12.

Table 5. Comparison of different sets of input features.

Input Features	Stable Samples	Unstable Samples	Overall Accuracy
	Accuracy	Accuracy	
V	0.9926	0.9038	0.9839
P_e	0.9947	0.9231	0.9877
$[V, P_e]$	0.9958	0.9327	0.9896

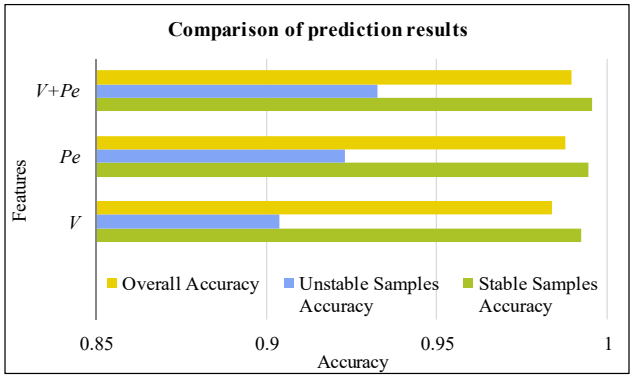


Figure 12. Comparison of prediction results with different features.

From Table 5, the prediction accuracy is only improved slightly from 98.77% to 98.98% by integrating both features P_e and V into one SVM prediction model. A more detailed comparison in Figure 12 shows that the third set of input features $[V, P_e]$ has the highest accuracy. However, the prediction accuracy for unstable samples (in blue) is still not satisfying; that is, the conservativeness of the prediction model is still not good enough. Therefore, it is necessary to find better solutions to reduce the error rate of unstable samples, which are more harmful to the operation of power systems.

3.4.2. Improved WTA Ensemble Learning Results for Conservative Prediction

For the two selected input features of V and P_e, the prediction accuracies of the two sub-learning machines and WTA ensemble learning model are shown in Table 6 and Figure 13.

Table 6. The prediction accuracy of different prediction models.

Input Features	Stable Samples	Unstable Samples	Overall Accuracy
	Prediction Accuracy	Prediction Accuracy	
V	0.9926	0.9038	0.9839
P_e	0.9947	0.9231	0.9877
V WTA P_e	0.9873	0.9926	0.9878

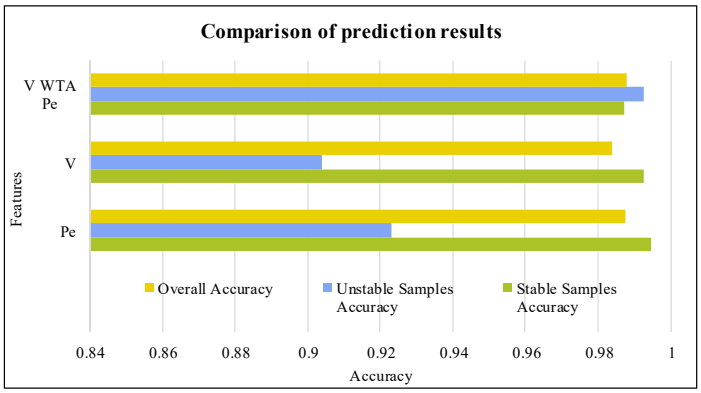


Figure 13. Comparison of accuracy in different models.

It can be seen from Table 6 and Figure 13 that the WTA model is able to improve the prediction accuracy of unstable samples (in blue) greatly, from 90.38%, 92.31% to 99.26%, while the overall prediction accuracy also increases slightly. Due to the special treatment to improve the model conservativeness for unstable situations, the accuracy of the proposed WTA ensemble model is higher than the recent work, such as the DSEC ensemble model of less than 97.03% in Reference [6]; the datasets were generated from the same IEEE 39-bus system. Therefore, the proposed WTA ensemble learning model can provide a strong basis for the online applications of TSA, based on machine learning technology.

4. Discussion

The WTA model is able to improve the prediction accuracy of unstable samples to 99.26%. In order to implement the proposed machine learning model into the online TSA, the higher the prediction accuracy for the unstable situations, the better. Any missed instability situation may lead to loss of synchronization, cascading failures, or even large-scale power outages. Making fast and accurate online transient stability predictions is not enough in transient operation of a power system appropriate online emergency control measures will be of interest in future research.

Another concern in practically implementing the proposed machine learning model into large-scale power systems is that the feature reduction should be further studied, because too many input features mean a large amount of measurement investment and a huge computational burden in real-time.

In addition, the main focus of this study is to perform machine learning-based transient stability predictions that only deal with the rotor angle stability of power systems. If appropriate data are available, the transient stability prediction modeling method based on machine learning can also be extended to small-disturbance stability, voltage stability, and frequency stability issues.

5. Conclusions

Recent artificial intelligence and machine learning technologies enable the use of online information of electrical and electromechanical conditions to be used to diagnose and predict the operating status of power systems. A high-accuracy conservative transient stability prediction model is proposed in this paper. Compared with the existing models, our model contains four improvements. (1) The sampled data containing multiple time stages are used as input features for the SVM classifier. It is found that the sampled data containing all three-time stages, namely the pre-fault, during-fault and post-fault, can better characterize the transient stability of the power systems; (2) the new feature of generators' electromagnetic power is found to be highly correlated to system stability. The SVM classification results show that the prediction accuracy of electromagnetic power feature is higher than the conventional generator rotor angle, generator rotor speed and bus voltage amplitude features; (3) electromagnetic power and voltage amplitude are determined as two superior features so as to reduce computational burden, based on mRMR feature selection; (4) a high-precision WTA ensemble

learning model based on the two selected features is established for power system transient stability prediction, which improves the accuracy for unstable situations from 90.38%, 92.31% to 99.26%. The WTA ensemble learning can significantly improve the conservativeness of the prediction model, and the overall prediction accuracy will also be increased slightly. All the research results are verified by the simulated samples on the IEEE 39-Bus system.

Author Contributions: Conceptualization, J.L. and H.S.; methodology, J.L.; software, H.S.; validation, H.S. and Y.L.; formal analysis, W.F.; data curation, S.N.; writing—original draft preparation, H.S.; writing—review and editing, J.L. and W.F.; supervision, W.F.; funding acquisition, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the National Natural Science Foundation of China under grant 51507126, in part by the Key Research and Development Program of Shaanxi under Grant 2017ZDCXL-GY-02-03, and in part by the Fundamental Research Funds for the Central Universities of China under Grant xjj2017145.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gan, G.; Zhu, Z.; Geng, G.; Jiang, Q. An efficient parallel sequential approach for transient stability emergency control of large scale power system. *IEEE Trans. Power Syst.* **2018**, *33*, 5854–5864. [CrossRef]
2. Ma, S.; Chen, C.; Liu, C.; Shen, Z. A measurement-simulation hybrid method for transient stability assessment and control based on the deviation energy. *Int. J. Electr. Power* **2020**, *115*, 105422. [CrossRef]
3. Rahman, M.S.; Pota, H.R.; Orchi, T.F. A multi-agent approach for enhancing transient stability of smart grids with renewable energy. *Int. J. Electr. Power* **2015**, *67*, 488–500. [CrossRef]
4. Zakariazadeh, A.; Jadid, S.; Siano, P. Stochastic operational scheduling of smart distribution system considering wind generation and demand response programs. *Int. J. Electr. Power* **2014**, *63*, 218–225. [CrossRef]
5. Liu, J.; Fang, W.; Zhang, X.; Yang, C. An improved photovoltaic power forecasting model with the assistance of aerosol index data. *IEEE Trans. Sustain. Energy* **2015**, *6*, 434–442. [CrossRef]
6. Chen, Z.; Han, X.; Fan, C.; He, Z.; Su, X.; Mei, S. A data segmentation-based ensemble classification method for power system transient stability status prediction with imbalanced data. *Appl. Sci.* **2019**, *9*, 4216. [CrossRef]
7. Gurung, S.; Naetiladdanon, S.; Sangswang, A. Coordination of power-system stabilizers and battery energy-storage system controllers to improve probabilistic small-signal stability considering integration of renewable-energy resources. *Appl. Sci.* **2019**, *9*, 1109. [CrossRef]
8. Xia, S.; Zhang, Q.; Hussain, S.T.; Hong, B.; Zou, W. Impacts of integration of wind farms on power system transient stability. *Appl. Sci.* **2018**, *8*, 1289. [CrossRef]
9. Sobbouhi, A.R.; Vahedi, A. Online synchronous generator out-of-step prediction by electrical power curve fitting. *IET Gener. Transmiss. Distrib.* **2020**, *14*, 1169–1176. [CrossRef]
10. Hazari, M.; Mannan, M.A.; Muyeen, S.M.; Umemura, A.; Takahashi, R.; Tamura, J. Stability augmentation of a grid-connected wind farm by fuzzy-logic-controlled DFIG-based wind turbines. *Appl. Sci.* **2018**, *8*, 20. [CrossRef]
11. Iravani, A.; de Leon, F. Real-time transient stability assessment using dynamic equivalents and nonlinear observers. *IEEE Trans. Power Syst.* **2020**. [CrossRef]
12. Liu, J.; Sun, H.; Wu, L.; Zhang, Z.; Niu, S.; Ke, X.; Huo, C. An overview of transient stability assessment of power systems. *Smart Power* **2019**, *47*, 44–53.
13. Karami, A.; Esmaili, S.Z. Transient stability assessment of power systems described with detailed models using neural networks. *Int. J. Electr. Power* **2013**, *45*, 279–292. [CrossRef]
14. Pai, M.A.; Vittal, V. Multimachine stability analysis using vector Lyapunov functions with inertial-centre decomposition. *Int. J. Electr. Power* **1983**, *5*, 139–144. [CrossRef]
15. Wang, B.; Sun, K.; Su, X. A decoupling based direct method for power system transient stability analysis. In Proceedings of the 2015 IEEE Power & Energy Society General Meeting, Denver, CO, USA, 26–30 July 2015; pp. 1–5.

16. Chiang, H.D.; Wu, F.F.; Varaiya, P.P. A BCU method for direct analysis of power system transient stability. *IEEE Trans. Power Syst.* **1994**, *9*, 1194–1208. [CrossRef]
17. Alinezhad, B.; Karegar, H.K. Out-of-step protection based on equal area criterion. *IEEE Trans. Power Syst.* **2017**, *32*, 968–977. [CrossRef]
18. Wehenkel, L.; van Cutsem, T.; Ribbens-Pavella, M. An artificial intelligence framework for online transient stability assessment of power systems. *IEEE Trans. Power Syst.* **1989**, *4*, 789–800. [CrossRef]
19. AL-Masri, A.N.; Ab Kadir MZ, A.; Hizam, H.; Mariun, N. A novel implementation for generator rotor angle stability prediction using an adaptive artificial neural network application for dynamic security assessment. *IEEE Trans. Power Syst.* **2013**, *28*, 2516–2525. [CrossRef]
20. Yan, R.; Geng, G.; Jiang, Q.; Li, Y. Fast transient stability batch assessment using cascaded convolutional neural networks. *IEEE Trans. Power Syst.* **2019**, *34*, 2802–2813. [CrossRef]
21. Shi, Z.; Yao, W.; Zeng, L.; Wen, J.; Fang, J.; Ai, X.; Wen, J. Convolutional neural network-based power system transient stability assessment and instability mode prediction. *Appl. Energy* **2020**, *263*, 114586. [CrossRef]
22. Tan, B.; Yang, J.; Tang, Y.; Jiang, S.; Xie, P.; Yuan, W. A deep imbalanced learning framework for transient stability assessment of power system. *IEEE Access* **2019**, *7*, 81759–81769. [CrossRef]
23. Wu, S.; Zheng, L.; Hu, W.; Yu, R.; Liu, B. Improved deep belief network and model interpretation method for power system transient stability assessment. *J. Mod. Power Syst. Clean Energy* **2020**, *8*, 27–37. [CrossRef]
24. Aliyan, E.; Aghamohammadi, M.; Kia, M.; Heidari, A.; Shafie-khah, M.; Catalão, J.P. Decision tree analysis to identify harmful contingencies and estimate blackout indices for predicting system vulnerability. *Electr. Power Syst. Res.* **2020**, *178*, 106036. [CrossRef]
25. Amraee, T.; Ranjbar, S. Transient instability prediction using decision tree technique. *IEEE Trans. Power Syst.* **2013**, *28*, 3028–3037. [CrossRef]
26. Wei, H.U.; Zongxiang, L.U.; Shuang, W.U.; Zhang, W.; Yu, D.O.N.G.; Rui, Y.U.; Baisi, L.I.U. Real-time transient stability assessment in power system based on improved SVM. *J. Mod. Power Syst. Clean Energy* **2019**, *7*, 26–37.
27. Gomez, F.R.; Rajapakse, A.D.; Annakkage, U.D.; Fernando, I.T. Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements. *IEEE Trans. Power Syst.* **2011**, *26*, 1474–1483. [CrossRef]
28. Ye, S.; Li, X.; Wang, X.; Qian, Q. Power system transient stability assessment based on adaboost and support vector machines. In Proceedings of the 2012 Asia-pacific Power & Energy Engineering Conference, Shanghai, China, 27–29 March 2012; pp. 1–5.
29. MA, Q.; Yang, Y.H.; Liu, W.Y.; Qi, Z.; Guo, J.Z. Power system transient stability prediction with combined SVM method mixing multiple input features. *Proc. CSEE* **2005**, *25*, 17–23. (In Chinese)
30. Wang, B.; Fang, B.; Wang, Y.; Liu, H.; Liu, Y. Power system transient stability assessment based on big data and the core vector machine. *IEEE Trans. Smart Grid* **2016**, *7*, 2561–2570. [CrossRef]
31. Li, Y.; Yang, Z. Application of EOS-ELM with binary Jaya-based feature selection to real-time transient stability assessment using PMU data. *IEEE Access* **2017**, *5*, 23092–23101. [CrossRef]
32. Lv, J.; Pawlak, M.; Annakkage, U.D. Prediction of the transient stability boundary using the lasso. *IEEE Trans. Power Syst.* **2013**, *28*, 281–288. [CrossRef]
33. Amjady, N.; Majedi, S.F. Transient stability prediction by a hybrid intelligent system. *IEEE Trans. Power Syst.* **2007**, *22*, 1275–1283. [CrossRef]
34. Kundur, P.; Paserba, J.; Ajjarapu, V.; Andersson, G.; Bose, A.; Canizares, C.; Hatziargyriou, N.; Hill, D.; Stankovic, A.; Taylor, C.; et al. Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions. *IEEE Trans. Power Syst.* **2004**, *19*, 1387–1401.
35. Chow, J.H.; Cheung, K.W. A toolbox for power system dynamics and control engineering education and research. *IEEE Trans. Power Syst.* **1992**, *7*, 1559–1564. [CrossRef]
36. Liu, J.; Wang, X.; Sun, H.; Cheng, L.; Ke, X.; Sun, X.; Wei, P. A conservative prediction model of power system transient stability. In Proceedings of the 2018 IEEE Power and Energy Society General Meeting, Portland, OR, USA, 5–10 August 2018; pp. 1–5.
37. Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal.* **2005**, *27*, 1226–1238. [CrossRef] [PubMed]

38. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intel. Syst. Technol.* **2007**, *2*, 1–27. [CrossRef]
39. Palm, R.B. *Prediction as a Candidate for Learning Deep Hierarchical Models of Data*; Technical University of Denmark: Kongens Lyngby, Denmark, 2012.