*mathematics*

*Article*

# Power System Transient Stability Assessment Based on Machine Learning Algorithms and Grid Topology

**Mihail Senyuk [1], Murodbek Safaraliev [1], Firuz Kamalov [2] and Hana Sulieman [3,*]**

1 Department of Automated Electrical Systems, Ural Federal University, 620002 Yekaterinburg, Russia
2 Department of Electrical Engineering, Canadian University Dubai, Dubai P.O. Box 415053, United Arab Emirates
3 Department of Mathematics and Statistics, American University of Sharjah, Sharjah P.O. Box 26666, United Arab Emirates
* Correspondence: hsulieman@aus.edu

**Abstract:** This work employs machine learning methods to develop and test a technique for dynamic stability analysis of the mathematical model of a power system. A distinctive feature of the proposed method is the absence of a priori parameters of the power system model. Thus, the adaptability of the dynamic stability assessment is achieved. The selected research topic relates to the issue of changing the structure and parameters of modern power systems. The key features of modern power systems include the following: decreased total inertia caused by integration of renewable sources energy, stricter requirements for emergency control accuracy, highly digitized operation and control of power systems, and high volumes of data that describe power system operation. Arranging emergency control in these new conditions is one of the prominent problems in modern power systems. In this study, the emergency control algorithms based on ensemble machine learning algorithms (XGBoost and Random Forest) were developed for a low-inertia power system. Transient stability of a power system was analyzed as the base function. Features of transmission line maintenance were used to increase accuracy of estimation. Algorithms were tested using the test power system IEEE39. In the case of the test sample, accuracy of instability classification for XGBoost was 91.5%, while that for Random Forest was 81.6%. The accuracy of algorithms increased by 10.9% and 1.5%, respectively, when the topology of the power system was taken into account.

**Keywords:** ensemble machine learning; extreme gradient boosting; power system modeling; random forest; transient stability

**MSC:** 68T01



**Citation:** Senyuk, M.; Safaraliev, M.; Kamalov, F.; Sulieman, H. Power System Transient Stability Assessment Based on Machine Learning Algorithms and Grid Topology. *Mathematics* **2023**, *11*, 525. https://doi.org/10.3390/math11030525

Academic Editors: Dongsheng Yu, Muhammad Junaid, Samson Yu and Yihua Hu

Received: 30 December 2022
Revised: 15 January 2023
Accepted: 16 January 2023
Published: 18 January 2023



## 1. Introduction

Application of algorithms based on machine learning (ML) in planning, operation, and control of power systems has become possible due to industry digitization and the collection of sufficient amounts of data. The following problems are solved using ML algorithms: equipment monitoring [1], load forecasting [2], forecasting of renewable sources of energy (RES) [3], adjustment of power system control devices [4], state estimation [5], and disturbance detection [6]. However, in modern operation and control of power systems, the preference is still given to the conventional methods based on deterministic approaches. On the other hand, ML methods have evolved into effective tools in analysis and control of power systems in terms of time response, accuracy, and adaptability. In addition, the ever-growing use of phasor measurements [7] will significantly increase the amount of data being exchanged between power system facilities and control centers, which has a considerable impact on the operation time of conventional methods [8] in terms of power system analysis.

One of the key tasks in controlling the electrical modes of power systems is assessment of the dynamic stability of synchronous generators with the subsequent application of emergency control. Evaluation of the dynamic stability of a synchronous generator is a complex nonlinear problem with a wide range of influencing factors: pre-emergency loading of synchronous generators, the place of occurrence of a short circuit, the duration of a short circuit, pre-emergency voltages in the nodes of the electrical network, etc. The high speed of transient processes, the nonlinear relationship between the parameters of the electrical regime and the large dimensions (up to several thousand nodes and branches) of the protected sections of power systems, and the presence of errors in the parameters of power equipment lead to significant time costs for assessing dynamic stability due to the use of methods with a rigidly defined algorithmic basis [9].

In conditions of digital transformation, increased data flows, and possible constraints of operation time for conventional algorithms, ML methods provide new opportunities. In particular, transient stability analysis can be a potential application of ML methods. The traditional methods of solving this problem are:

- Numerical integration of a system of algebraic differential equations, which describe the dynamic behavior of a power system;
- Energy-based analysis (equal area criterion) [9].

These approaches have drawbacks related to the significant amount of time required for numerical integration of large power system models and possible insufficient accuracy of energy-based methods. ML methods can be applied to negate these drawbacks, which could considerably reduce the processing time while maintaining required accuracy.

The use of ML methods makes it possible to significantly speed up the process of assessing dynamic stability due to the absence of the operation of numerical differentiation of a nonlinear system of algebraic differential equations describing a dynamic model of the power system. In addition, the use of ML methods makes it possible, in the training, to take into account sample data that is obtained not only during mathematical modeling of transients on a simulation model but also from measurements of real transients observed during disturbances occurring in real power systems. The advantages of ML methods described above allow us to move to a qualitatively new level of dynamic stability analysis due to greater adaptability and consideration of data of different natures: mathematical and physical.

The purpose of this paper is to assess the impact on the accuracy of ML methods of topological connectivity of the electrical network and to develop recommendations for training and using ML algorithms to assess dynamic stability.

## 2. Literature Review

The problem of transient stability analysis of power systems is considered to be non-linear and large scale. Traditionally, dynamic response of a power system is found using numerical differentiation of algebraic differential equations, which describe dynamic models of synchronous generators, loads, transmission lines, emergency control devices, etc. The results of this numerical differentiation greatly depend on parameters of models in use and are not without considerable time costs. Time costs of small signal stability analysis can be reduced by means of linearization of the equation system that describes the behavior of a power system. Linearization cannot be applied to the problem of transient stability due to rapid and significant changes of electrical parameters during transients.

To minimize the time cost of numerical differentiation, qualitative methods based on analysis of kinetic energy of generator rotor during a disturbance and potential energy in post-disturbance state are used. The number of calculations can be drastically decreased by utilizing reduced models of power systems [10].

The next stage in the improvement of accuracy and operation time of transient stability analysis of power systems has been the application of highly accurate PMUs [11]. These devices have made it possible to significantly increase the sampling rate of power system parameter estimation, thus becoming an effective tool of big data analysis in power systems in real time.

The growing amount of digital data and the continually developing theory of ML have led to the emergence of completely new methods of transient stability analysis. The input data can be represented by measurements obtained from instrument transformers on facilities of a power system and results obtained from simulations of processes in the power system under consideration. These sources of data have their own advantages and drawbacks. The measurement data are more informative in terms of the power system state, considering various control systems with their settings. Digital simulations of a power system can be used in analysis of a set of power system states for the purposes of emergency and dispatch control. Both data sources are regularly used to form learning and testing data sets for the ML-based methods of transient stability analysis.

In the framework of the ML theory, the problem of transient stability analysis can be reduced to the problem of binary classification with two classes: unstable state and stable state. Since the unstable state is much rarer than the stable one, the sample data become unbalanced, which has a huge impact on learning and operation of an ML model. When learning and testing samples are formed, the data may be noisy and incomplete and possess outliers.

Finding the features that have a significant impact on the state of the class (stability) is of great importance in forming the testing and the learning samples. The following signals of power system parameters can be considered in order to analyze transient stability: angular velocity of generator rotor, load angle, active and reactive power, etc.

Features of the frequency domain can be considered as well. Many conventional ML methods are not designed to operate with signals from the time domain. This obstacle can be overcome by extracting two specific points from the time domain, namely before and after a disturbance. This procedure reduces the times series to two points [12], thereby simplifying the problem to be solved and increasing the speed of learning and that of the ML model operation. The further reduction of dimension number can be accomplished via principal components analysis [13], singular value decomposition, or linear discriminant analysis. The use of the method of random selection of features for the purpose of learning time minimization was suggested in [14].

Since loss of transient stability is a relatively rare event in modern power systems, it is important to construct a dataset by means of performing series of electromechanical transient calculations using power system models. These calculations are carried out with varying load and generation at nodes of a power system in addition to varying location, type, and duration of faults [15]. The selection from the obtained dataset is made on the basis of the distribution recommended in [16], where 10% are three-phase faults, 20% are two-phase faults, and 70% are single-phase faults. The resultant sample is unbalanced most of the time, with stable states being prevalent. The stratified random sampling is conducted to separate learning and testing samples while maintaining data balance.

The problem of transient stability analysis of power systems can be solved using the following ML algorithms:

- Artificial neural networks (ANN);
- Support vector machine (SVM);
- Random forest (RF);
- Ensemble algorithms (EA);
- Deep learning (DL).

One of the first studies that covered ANN application in transient stability analysis described ANNs with a single hidden layer [17] or parallel ANNs. However, due to lack of computational capacity, ANNs have not found widespread use. As computational capabilities of equipment began to grow, the deep learning algorithms have become a tool for the transient stability analysis [18–20]. These algorithms are based on ANNs with multiple layers of different types. The main difficulty of applying the deep learning is related to initialization of layers, selection of the activation function, finding the optimal learning speed, etc. In the framework of transient stability analysis, feature selection is not assumed to be performed in the deep learning algorithms, unlike in other ones.

Transient stability analysis can be reduced to the problem of binary classification with two classes: stable state or unstable state. Hence, classification methods are often used to solve the problem. SVM is one of the major algorithms [21–23]. Advantages of using the SVM include short time response and high reliability caused by quadratic programming procedures in convex regions with a single solution. The main drawbacks are high sensitivity to noise in input data and absence of a common approach to automatic algorithm core selection in the case of linear partibility of classes.

The RF algorithm is based on using the aggregations of elementary classification algorithms, called decision trees. The classification result for each separate tree is obtained with a weight, which is a priori assigned during learning. In [24–26], the RF algorithm was used to evaluate transient stability of complex power systems. Advantages of this algorithm include high efficiency due to the possibility of using parallel calculations and a small number of parameters, leading to simplicity of adjustment. The main disadvantage is its proneness to re-learning.

The ensemble algorithm is presented by multiple basic algorithms. The common solution is obtained by means of solving equations for each basic sub-algorithm. The subsets of bagging, boosting, and stacking compose the ensemble algorithms. The boosting and the bagging algorithms use results of basic algorithms, while the stacking algorithm uses the second-level model, which learns using data from the basic models. The ensemble algorithms were used in transient stability analysis of power systems in [27–29]. Advantages of the EA are the following: great generalization ability, ability to identify outlying data, and short time response. Drawbacks include proneness to re-learning and necessity to form relatively large learning samples.

The XGBoost is considered to be one of the most effective EAs [30]. There are several reasons for choosing this algorithm as the tool of transient stability analysis of power systems: ability to process data that have gaps and noise; almost unsurpassed ability to work with table data in solving problems of classification and regression; and high win rate in tests at kaggle.com (accessed on 3 January 2023). In addition, the XGBoost algorithm has high accuracy and efficiency, re-learning-based tuning via regularization, and a large number of adjustable parameters.

Other ML algorithms are also applied in solving the problem of transient stability analysis. The gradient boosting was used in [31], while the kernel regression was used in [32].

In the study [33], the DL algorithm was used to assess the dynamic stability of complex power systems. The algorithm was tested on real data recorded during the historical observation of disturbances occurring in the power system. A parallel convolution algorithm was used to eliminate redundant input features. The assessment of the dynamic stability of the power system was carried out on the basis of the Deep Forest algorithm, in which the authors made a number of improvements. The test result showed the advantages of the improved algorithm used in forecasting accuracy, learning speed, and update time.

DL is able to generate new functions from a limited set of functions located in the training data set. Therefore, deep learning algorithms can create new tasks to solve the current ones. Another advantage of DL algorithms relates to the ability to determine the most important functions, which allows deep learning to effectively provide specialists with an accurate and reliable analysis result. The main drawback of DL algorithms is associated with the complexity of interpreting the results obtained and the high resource intensity.

The Table 1 shows a comparison of ML algorithms.

**Table 1.** Comparison of ML algorithms.

| Algorithm | Advantages | Disadvantages |
|---|---|---|
| ANN [17] | Short time response | Time-consuming learning procedures |
| SVM [21–23] | Short time response | Sensitivity to noise in the input data, no automatic selection of the algorithm core |
| RF [24–26] | Parallel calculations, simple adjustment | Proneness to re-learning |
| EA [27–29] | Short time response | Large amounts of data are required |
| XGBoost [30] | Short time response, high accuracy | Large amounts of data are required |
| DL [33] | Ability to create new functions from a limited set of functions, the ability to identify the most important functions | Complexity of interpretation of the results obtained, high resource intensity |

The key requirements for an ML algorithm for transient stability analysis of power systems are efficiency and time response. On the basis of these conditions, XGBoost and RF were selected as the algorithms in this study.

Composition of the learning sample is crucial since it directly affects the results of transient stability analysis. Generally, transient stability depends on the following parameters:

- Pre-disturbance active power output of a synchronous generator;
- Voltage at generator bus;
- Impedance between fault location and generator bus;
- Governor control response.

Influence of power system topology on transient stability analysis was not considered in the reviewed studies. The total impedance between fault location and generator bus depends on fault type and power system topology, i.e., operating states of transmission lines and transformers. Hence, the purpose of this research is to study of topology influence on transient stability analysis on the basis of ML methods.

The original contribution of this paper is to consider the topological connection of the electrical network when identifying the loss of dynamic stability of a synchronous generator on the basis of machine learning methods. This task is particularly relevant because of the obvious influence of the total resistance from the synchronous generator in question on the short-circuit point. Another important task is to determine the degree of influence of the topology of the electrical network on the accuracy of identification of the loss of dynamic stability of the synchronous generator.

## 3. Transient Stability Analysis of a Test Power System Model

The test studies were carried out through simulations of electromechanical transients using Matlab/Simulink. The machine learning algorithms were implemented using the Python3 library called Scikit-learn.

### *3.1. Data Sampling*

Simulations were carried out on the test power system IEEE39 (Figure 1). These results were used in sampling of learning and testing data for algorithms XGBoost and Random Forest. A description of the model and its parameters can be found in [34].

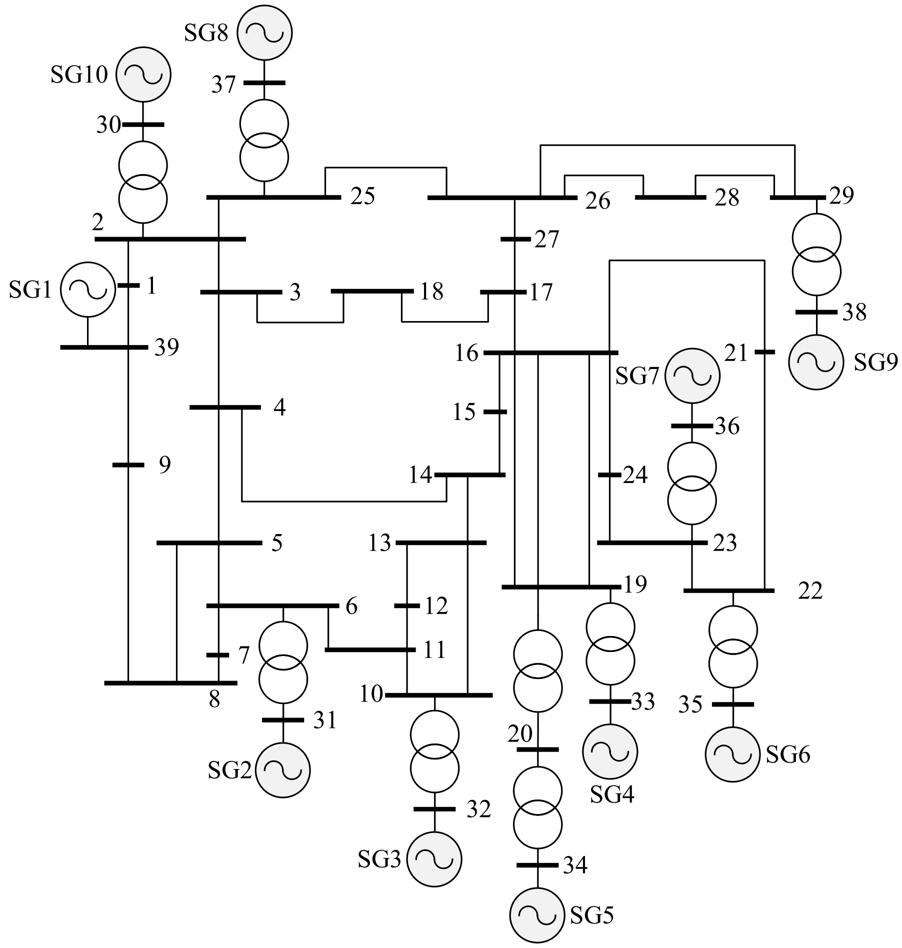


**Figure 1.** Test power system IEEE39.

A series of electromechanical transients were simulated to form sampling sets with the following parameters:

- Power outputs of synchronous generators (nine generators in total; Generator 1 is performing the role of the external power system, and its stability is not taken into account) were varied in the range of 60%, 80%, and 100% of their rated capacity;
- Fault locations (buses): Single-phase, two-phase, and three-phase faults with duration of 0.15 s were simulated in each bus (39 buses in total);
- Topology of the power system: Maintenance of a single line was simulated (37 lines in total), no transformer maintenance was considered.

Thus, the total amount of electromechanical transient simulations ($N$) is:

$$N = N_1 \cdot N_2 \cdot N_3 \cdot N_4 = 3 \cdot 3 \cdot 38 \cdot 37 = 12{,}654, \quad (1)$$

where $N_1$—number of generator power output options, $N_2$—number of fault types, $N_3$—number of fault locations–buses (no fault is simulated in Bus 39), and $N_4$—number of single line maintenance events.

Occurrence of transient instability is detected by load angle of the generator ($\delta$) exceeding 360 degrees. Figure 2 illustrates three-phase fault at Bus 23, which results in instability of Generators SG6 and SG7.

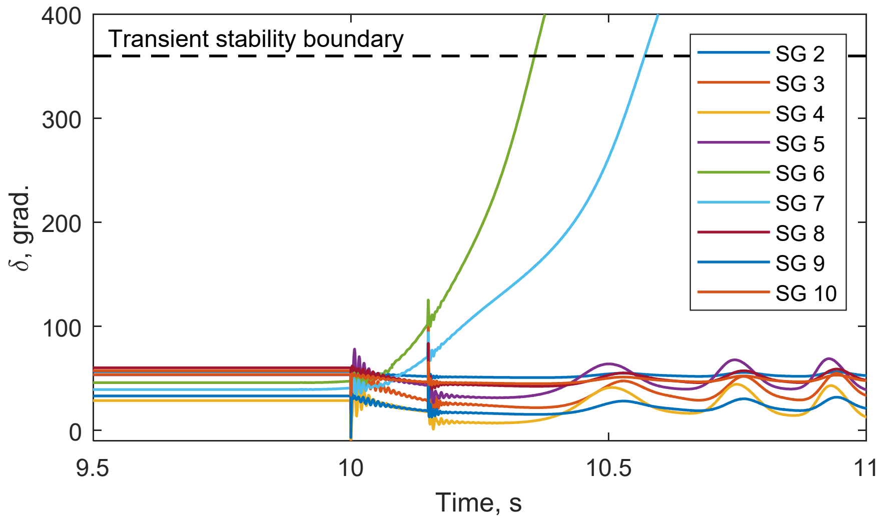


**Figure 2.** Example of a transient leading to instability of Generators SG6 and SG7.

The data sample was obtained as the result of a series of electromechanical transients (Table 2).

**Table 2.** Structure of the obtained data sample.

| No | Name | Description |
|---|---|---|
| | | Features |
| 1 | f1 | Maintenance of Line 1–2 (values: 1 or 0) |
| 2 | f2 | Maintenance of Line 2–3 (values: 1 or 0) |
| 37 | f37 | Maintenance of Line 39–1 (values: 1 or 0) |
| 38 | f38 | Active power outputs of Generators SG 2–10 (values: 0.6, 0.8, or 1) |
| 39 | f39 | Fault at Bus 2 (values: 1 or 0) |
| 77 | f77 | Fault at Bus 38 (values: 1 or 0) |
| 78 | f78 | Single-phase fault (values: 1 or 0) |
| 79 | f79 | Two-phase fault (values: 1 or 0) |
| 80 | f80 | Three-phase fault (values: 1 or 0) |
| 81 | f81 | Pre-disturbance load angle of Generator SG 2 (values: from 0 to 180) |
| 90 | f90 | Pre-disturbance load angle of Generator SG 10 (values: from 0 to 180) |
| 91 | f91 | Pre-disturbance voltage of stator winding of Generator SG 2 (values: from 0.5 to 1.2 from rated voltage) |
| 100 | f100 | Pre-disturbance voltage of stator winding of Generator SG 10 (values: from 0.5 to 1.2 from rated voltage) |
| | | Target |
| 1 | t | Transient stability of Generators SG 2–10 (values: 1 or 0) |

Since the sample was obtained as a result of mathematical simulation, it has no gaps, outliers, or noise.

The issue of the occurrence of fluctuations in the post-emergency mode is very important. The nature of such fluctuations is related to the settings of the system regulators. In the algorithms considered, the influence of the control systems is considered directly by taking into account the load angle of the synchronous generator and the voltage on the stator winding in the training and test samples.

The Spearman correlation was analyzed to assess impact of each feature on the results of classification:

$$\text{SpearmanC} = 1 - \frac{6\int d^2}{n(n^2-1)}, \tag{2}$$

where SpearmanC—Spearman correlation value, $d$—difference between ranks for the pair (X,Y) of two numbers series, and $n$—length of series X and Y. A rank is the number of a value in the series after sorting by value. The Spearman correlation can be applied to both numerical data and categories expressed numerically. The correlation can indicate the interdependence of varied parameters.

The results of features with ratio of correlation by absolute classification result of less than 10% are shown on Figure 3.

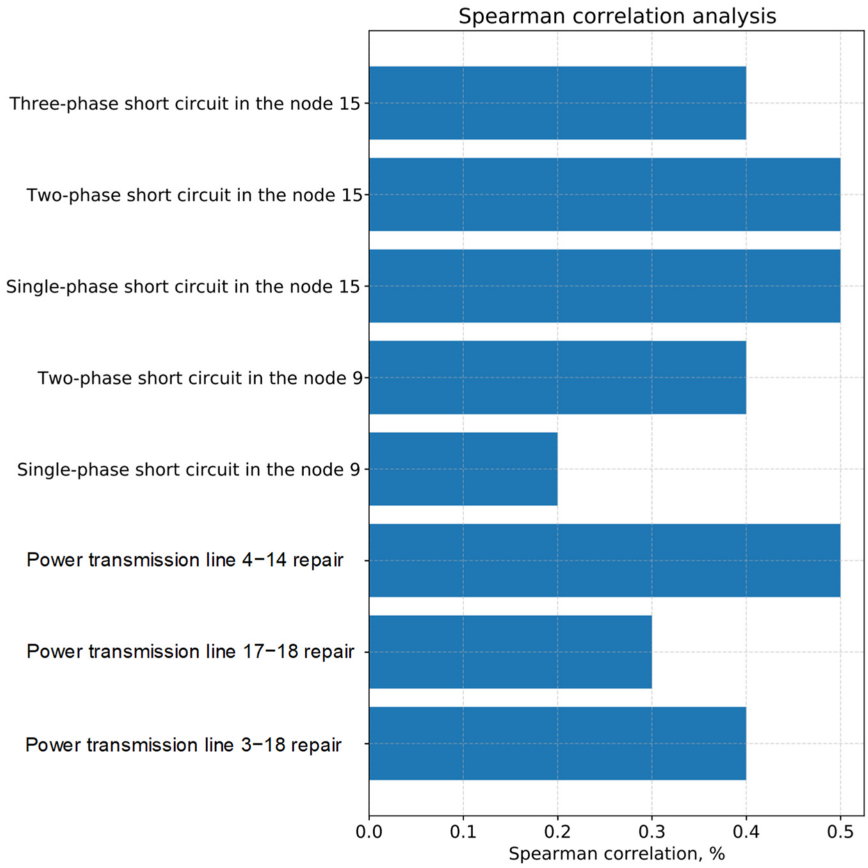


**Figure 3.** The results of data correlation analysis.

The Spearman correlation was calculated using the scipy library. Since the values of correlation of features in Figure 3 are considered insignificant, these features are removed from the data sample. The sample was separated according to the ratio of 80%:20% (80%—learning sample, 20%—testing sample).

### *3.2. The XGBoost Algorithm Learning Results*

The parameters were set using the function GridSearch in order to increase the accuracy of the XGB algorithm. This function is used to search through combinations of parameters with the goal of finding the combination that provides the greatest model quality metric. In the process of learning the XGBoost algorithm, the following parameters were obtained:

- $L_1$ regularization, penalty for weight functions (base value of 0): 0.1;
- $L_2$ regularization, penalty for weight functions (base value of 0): 0.2;
- The required minimum decrease of the function in a process of creating a new leaf (base value 0): 0.5;
- The maximum tree depth of the base classification system. This parameter determines complexity of the model and level of retrain (base value 3): 5;
- Base value of probability that a data line corresponds to binary cases; it allows for correction of the dropping accuracy due to class imbalance (base value 0.5): 0.8;
- A basic classificatory number of the composition; it controls complexity of the model: 400;
- Pace of gradient descent; it controls the possibility of losing the local minima: 1;

- Training sample rate; this parameter is selected randomly for training of one tree (base value 1): 0.8.

Feature importance for the trained model is described in Figure 4. Feature importance for the algorithm can be described by increased accuracy after using a feature in tree branches during splitting. Importance provides an estimate of how useful or valuable each feature of the sample is in constructing extended decision trees in the model. The more a feature is used to make key decisions with decision trees, the higher its relative importance. The importance is calculated explicitly for each feature in the dataset, which allows one to rank the features and compare them. The importance is calculated for one decision tree as the amount by which each feature separation point improves the productivity indicator, weighted by the number of observations for which the node is responsible. The measure of efficiency may be the Gini index used to select the separation points, or it may be another more specific error function.

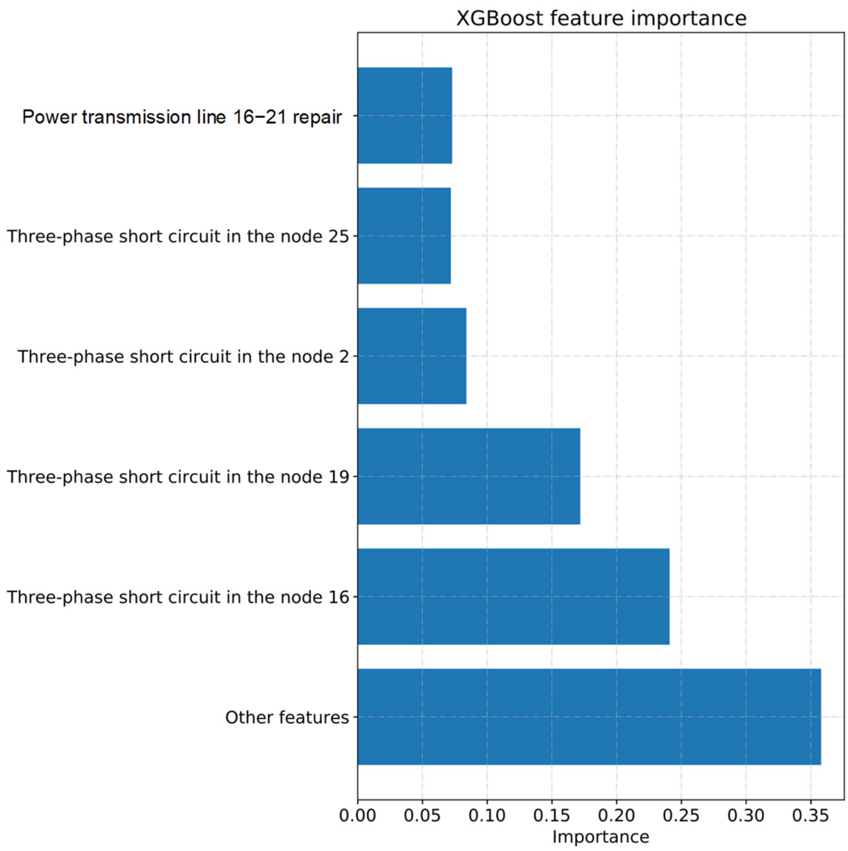


**Figure 4.** XGBoost feature importance.

The greatest importance for transient stability analysis of the test power system using XGBoost algorithm is the feature of three-phase fault at Bus 16 because this bus has a high number of connections with other buses. Figure 5 describes an example of one of the base classifiers obtained after training the XGBoost algorithm.

The rules of classification and the values in the algorithm's leaves are defined after the XGBoost algorithm learning. In test data processing, all the base classifiers were involved in the test data processing, which are the base for transient stability analysis.

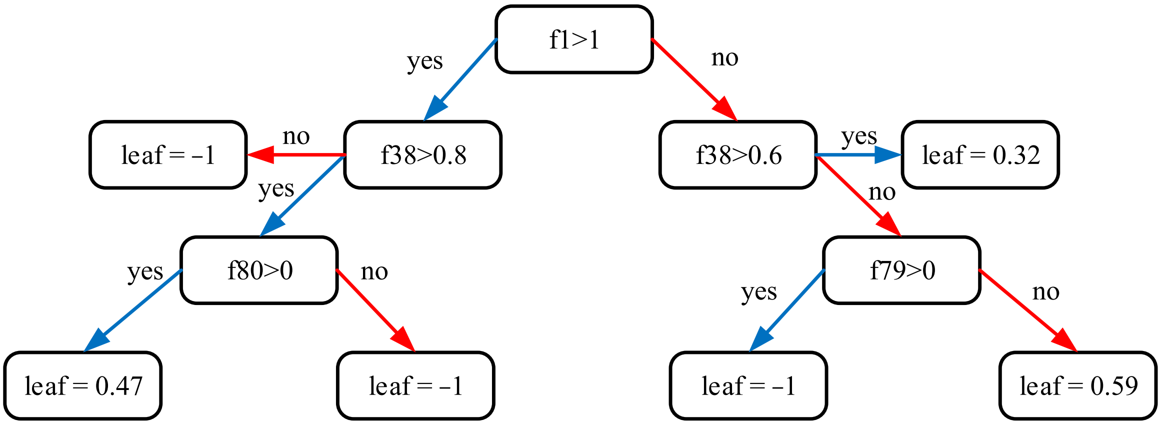


**Figure 5.** An example of one of the base classifiers of the XGBoost algorithm.

### *3.3. The Results of the Random Forest Algorithm Learning*

The parameters of RF were set using the function GridSearch to increase the algorithm accuracy. During the learning stage of the Random Forest algorithm, the following parameters were obtained:

- The number of base classifiers (base value 100): 550;
- Tree depth of a base classifier (base value 1): 3;
- The minimum number of copies of the data required for splitting (base value 2 strings): 0.005;
- The minimum rate of data copies in the leaf (base value 1 string): 0.002;
- Fraction of features in the training sample, chosen randomly for learning of one tree: 0.6;
- Class weights in the importance-graph-described weights are the following: for Class 0, where weight for the Class 1 is constant (base value 1 for every class): 6.

The feature importance values for the trained model are shown in Figure 6. The largest importance for instability classification is the feature of the three-phase fault at Bus 16. A similar result was obtained for the XGBoost algorithm.

A few basic classifiers of the trained Random Forest algorithm are shown in Figure 7. The information regarding a feature and the splitting rule is described in leaves or nodes of the tree; a number of data strings in a node expressed in %; and classification of a node or leaf in terms of class (0—transient instability, 1—transient stability).

The sequence of calculations to determine the importance of the Random Forest algorithm features is as follows:

- Selection of a random data set whose target variable is categorical;
- Dividing the data set into training and test parts;
- Calculation of the impurity node of each specific column where it branches. This is determined by calculating the right impurity and the left impurity branching off from the main node;
- Calculating the importance of the column function for this particular decision tree by calculating weighted averages of impurity nodes;
- The obtained values of the importance of the features will be averaged over the number of decision trees constructed. These obtained values of the importance of the features will be the final values in relation to the random forest classifier algorithm;
- The values will be in the range from 0 to 1. This will give a clearer idea of the choice of functions or columns for effective training of the model.

The rules of data splitting in every node of the base classifier are defined in the process of the Random Forest algorithm learning.

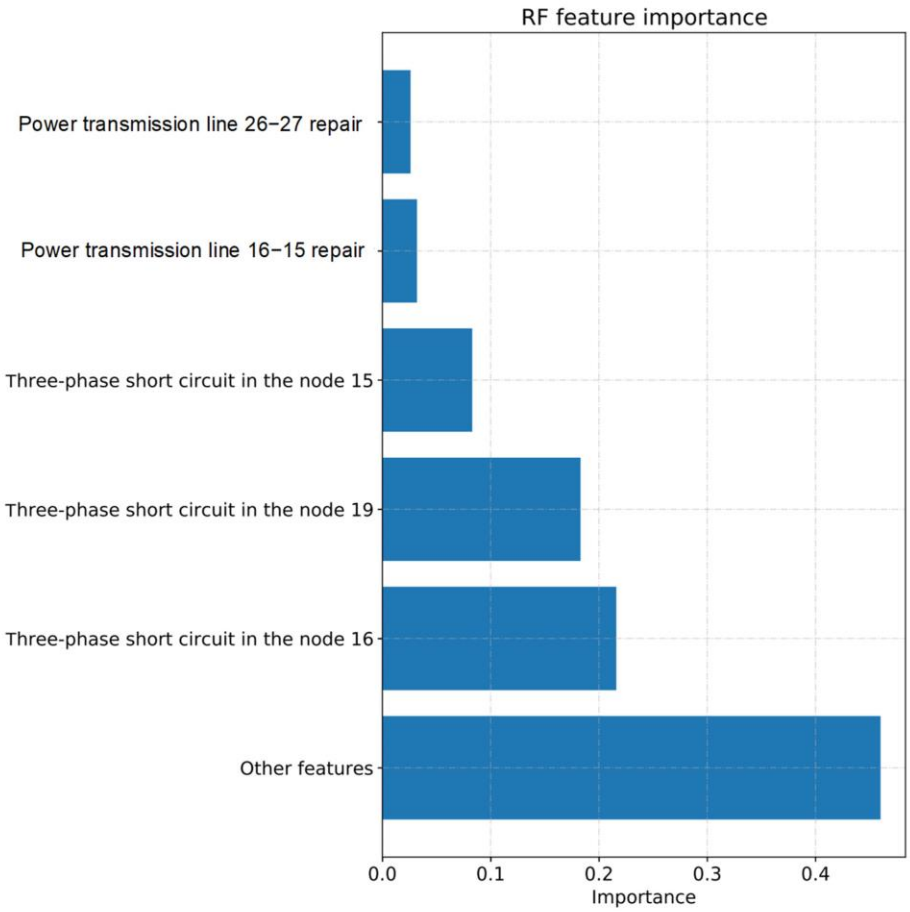


**Figure 6.** Feature importance for the Random Forest algorithm.

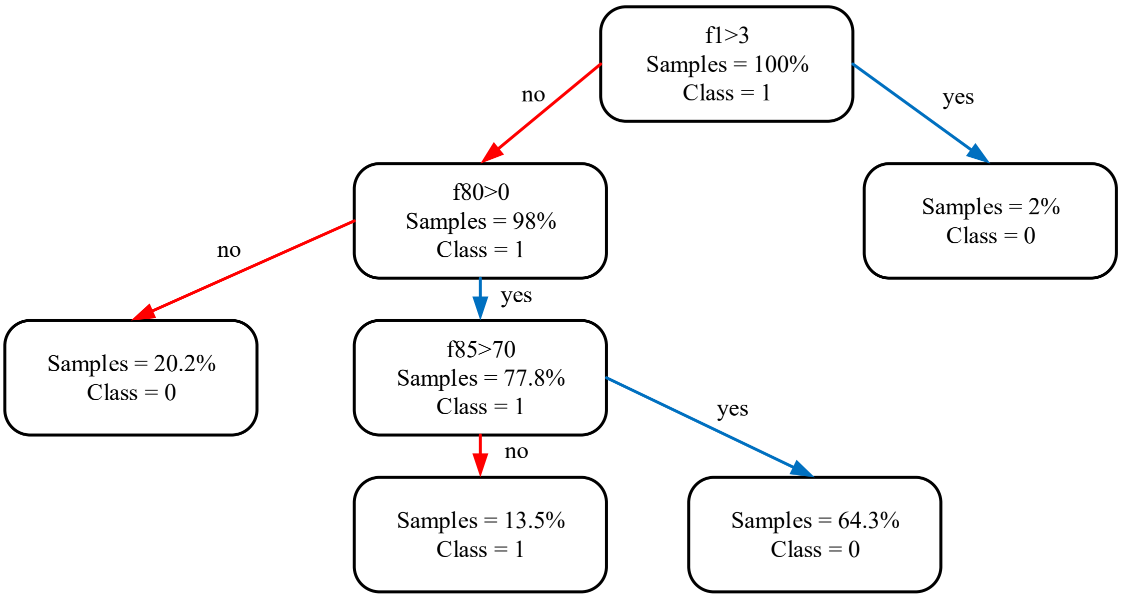


**Figure 7.** An example of one of the base classifiers of the Random Forest algorithm.

## 4. A Configuration of the Information-Gathering System for the Transient Stability Analysis of a Power System Based on the Machine Learning

Figure 8 describes the possible integration of the ML model for transient stability analysis.

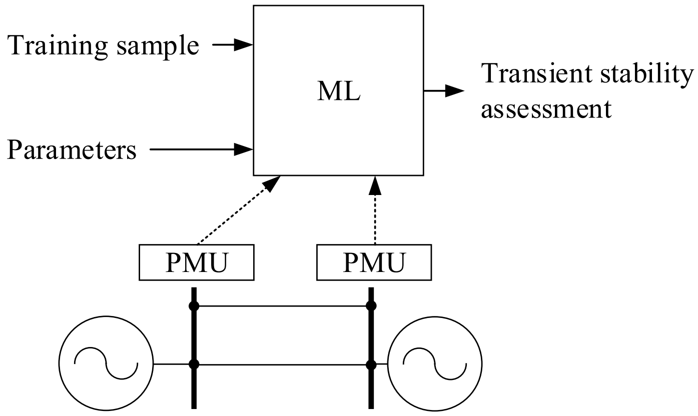


**Figure 8.** Possible ML model integration for transient stability analysis.

The results of mathematical modeling or data obtained from a real power system can be used as the learning set. In the case of using real measurements, it is necessary to filter the data, eliminate the outliers, and fill the gaps in the data.

To form a training sample, an approach of combining data obtained from the following sources can be used:

- Calculation results of electromechanical transients obtained from a verified model of the power system. In the practice of managing the modes of power systems, simulation models are used, the parameters of which can be taken from the passports of power equipment or test data. The values of loads, voltage levels, and other parameters of the electrical regime are obtained as a result of the assessment of the condition;
- Data obtained as a result of recordings of real transients occurring in the power system.

Such an approach to the formation of a training sample will ensure the sufficiency of its volume (through the use of mathematical modeling) and better representativeness due to the consideration of real processes occurring in the energy system.

After forming the learning sample, it is necessary to adjust the algorithm, which can be accomplished automatically or manually. After training and testing, the algorithm is ready to be used in the transient stability analysis of a power system.

It is suggested to use PMU data as the pre-disturbance data. Combinations of highly accurate PMU measurements and state estimation procedures [35] make it possible to obtain a complete list of state and equipment parameters, influencing transient stability. Results of transient stability analysis can be used in emergency control of power systems.

## 5. Results and Discussion

In this study, application of two ML algorithms for classifying the dynamic stability of the post-emergency operation of the power system was considered. XGBoost and Random Forest were chosen as the ML algorithms due to their effectiveness and the availability of anti-retraining techniques.

The comparative results of testing the XGBoost and Random Forest algorithms are presented in Table 3. The performance of the algorithms was based on the precision and recall metrics given by the following equations:

$$\text{precision} = \frac{tp}{tp + fp}, \tag{3}$$

$$\text{recall} = \frac{tp}{tp + fn}, \tag{4}$$

where *tp*—true positive decision, *fp*—false positive decision, and *fn*—false negative decision.

**Table 3.** Comparison of results of testing the trained algorithms.

| Parameter | XGBoost | Random Forest |
|---|---|---|
| | Accounting for power system topology | |
| Accuracy | 0.915 | 0.816 |
| Average accuracy between classes | 0.864 | 0.744 |
| Precision | 0.898 | 0.847 |
| Recall | 0.858 | 0.851 |
| | Not accounting for power system topology | |
| Accuracy | 0.806 | 0.801 |
| Average accuracy between classes | 0.817 | 0.711 |
| Precision | 0.815 | 0.816 |
| Recall | 0.824 | 0.801 |

It was found that the inclusion of the electrical network topological connectivity provides an increase in the algorithms' accuracy. In the case of the test sample, accuracy of the instability classification for XGBoost is 91.5%, while that for Random Forest is 81.6%. The consideration of the topology of a power system increases the accuracy of algorithms by 10.9% and 1.5%, respectively.

Thus, on the basis of the presented results, it can be concluded that the XGBoost algorithm is preferable to the Random Forest algorithm due to greater accuracy in classifying the loss of stability of power systems.

## 6. Conclusions

The algorithm of transient stability analysis of a power system based on algorithms XGBoost and Random Forest is presented in this study. The topology of a power system was taken into account. The literature review covers current studies of transient stability analysis based on machine learning methods.

The test power system was IEEE39. The data sample of 12,654 electromechanical transients was obtained as a result of power system simulations.

After learning, parameters of the algorithms XGBoost and Random Forest were obtained. In the case of the test sample, the accuracy of instability classification for XGBoost is 91.5% taking into account topology of the power system, while that for Random Forest is 81.6%. When topology is accounted for, the accuracy of the algorithms increases.

The practical application of the algorithms considered for system operators of the power system can be considered in the following tasks: monitoring of the reserve of dynamic stability and adaptive emergency management [36].

Further studies will cover development of the AI-based emergency control of a power system. This problem is drastically different from that of transient stability due to increased variability caused by selection of the optimal control actions set. In addition, the second direction of further research is the approbation of the considered machine learning algorithms on more complex models of power systems and on data obtained from real power systems.

Despite the proven effectiveness of using ML algorithms to assess the dynamic stability of power systems, the proposed method does not consider the presence or absence of low-frequency oscillations in the post-accident mode of operation. Consideration of this problem is planned in subsequent studies.

Another direction for further work is the approbation of the proposed method for assessing dynamic stability on data obtained from the real power system.

**Author Contributions:** Conceptualization, M.S. (Mihail Senyuk) and H.S.; methodology, M.S. (Murodbek Safaralie); software, M.S. (Mihail Senyuk); validation, F.K.; formal analysis, M.S. (Murodbek Safaralie), F.K. and H.S.; investigation, M.S. (Mihail Senyuk) and M.S. (Murodbek Safaralie); resources, F.K.; data curation, F.K.; writing—original draft, M.S. (Mihail Senyuk); writing—review and editing, M.S. (Murodbek Safaralie) and H.S.; supervision, M.S. (Murodbek Safaralie); project administration, H.S.; funding acquisition, H.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported, in part, by the Open Access Program from the American University of Sharjah.

**Data Availability Statement:** Not applicable.

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

## References

1. Khalyasmaa, A.I.; Senyuk, M.; Eroshenko, S.A. High-Voltage Circuit Breakers Technical State Patterns Recognition Based on Machine Learning Methods. *IEEE Trans. Power Deliv.* **2019**, *34*, 1747–1756. [CrossRef]
2. Wang, H.; Huang, X.; Gao, S.; Yang, Z.; Gao, T.; Zhao, Q.; Ding, H. Electric vehicle charging load clustering and load forecasting based on long short term memory neural network. In Proceedings of the 2022 IEEE 5th International Electrical and Energy Conference (CIEEC), Nangjing, China, 27–29 May 2022; pp. 3196–3200.
3. Keerthisinghe, C.; Mickelson, E.; Kirschen, D.; Shih, N.; Gibson, S. Improved PV Forecasts for Capacity Firming. *IEEE Access* **2020**, *8*, 152173–152182. [CrossRef]
4. Gelli, R.; Manimaran, G. Anomaly Detection and Mitigation for Wide-Area Damping Control using Machine Learning. In Proceedings of the 2020 IEEE Power & Energy Society General Meeting (PESGM), Montreal, QC, Canada, 2–6 August 2020.
5. Hong, G.; Kim, Y.S. Supervised Learning Approach for State Estimation of Unmeasured Points of Distribution Network. *IEEE Access* **2020**, *8*, 113918–113931. [CrossRef]
6. Dhibi, K.; Fezai, R.; Mansouri, M.; Trabelsi, M.; Kouadri, A.; Bouzara, K.; Nounou, H.; Nounou, M. Reduced Kernel Random Forest Technique for Fault Detection and Classification in Grid-Tied PV Systems. *IEEE J. Photovolt.* **2020**, *10*, 1864–1871. [CrossRef]
7. De Oliveira-De Jesus, P.M.; Rodriguez, N.A.; Celeita, D.F.; Ramos, G.A. PMU-Based System State Estimation for Multigrounded Distribution Systems. *IEEE Trans. Power Syst.* **2021**, *36*, 1071–1081. [CrossRef]
8. Hong, T.; Chen, C.; Huang, J.; Lu, N.; Xie, L.; Zareipour, H. Guest Editorial Big Data Analytics for Grid Modernization. *IEEE Trans. Smart Grid* **2016**, *7*, 2395–2396. [CrossRef]
9. Kundur, P. *Power System Stability and Control*; McGraw-Hill, Inc.: New York, NY, USA, 1994.
10. Glavic, M.; Ernst, D.; Ruiz-Vega, D.; Wehenkel, L.; Pavella, M. E-SIME—A method for transient stability closed-loop emergency control: Achievements and prospects. In Proceedings of the 2007 iREP Symposium—Bulk Power System Dynamics and Control-VII. Revitalizing Operational Reliability, Charleston, SC, USA, 19–24 August 2007; pp. 1–10.
11. Beryozkina, S.; Senyuk, M.; Berdin, A.; Dmitrieva, A.; Dmitriev, S.; Erokhin, P. The Accelerate Estimation Method of Power System Parameters in Static and Dynamic Processes. *IEEE Access* **2022**, *10*, 61522–61529. [CrossRef]
12. Dabou, R.T.; Kamwa, I.; Chung, C.; Mugombozi, C.F. Time Series-Analysis Based Engineering of High-Dimensional Wide-Area Stability Indices for Machine Learning. *IEEE Access* **2021**, *9*, 104927–104939. [CrossRef]
13. Chen, K. Indirect PCA Dimensionality Reduction Based Machine Learning Algorithms for Power System Transient Stability Assessment. In Proceedings of the 2019 IEEE Innovative Smart Grid Technologies—Asia (ISGT Asia), Chengdu, China, 21–24 May 2019; pp. 4175–4179.
14. Mi, D.; Wang, T.; Gao, M.; Li, C.; Wang, Z. Online transient stability margin prediction of power systems with wind farms using ensemble regression trees. *Int. Trans. Electr. Energ. Syst.* **2021**, *31*, e13057. [CrossRef]
15. Liu, T.; Liu, Y.; Xu, L.; Liu, J.; Mitra, J.; Tian, Y. Non-parametric statistics-based predictor enabling online transient stability assessment. *IET Gener. Transm. Distrib.* **2018**, *12*, 5761–5769. [CrossRef]
16. Sarajcev, P.; Kunac, A.; Petrovic, G.; Despalatovic, M. Power System Transient Stability Assessment Using Stacked Autoencoder and Voting Ensemble. *Energies* **2021**, *14*, 3148. [CrossRef]
17. Amjady, N.; Ehsan, M. Transient stability assessment of power systems by a new estimating neural network. *Can. J. Electr. Comput. Eng.* **1997**, *22*, 131–137. [CrossRef]
18. Hou, J.; Xie, C.; Wang, T.; Yu, Z.; Lü, Y.; Dai, H. Power System Transient Stability Assessment Based on Voltage Phasor and Convolution Neural Network. In Proceedings of the 2018 IEEE International Conference on Energy Internet (ICEI), Beijing, China, 21–25 May 2018; pp. 247–251.
19. Zhou, Y.; Zhao, W.; Guo, Q.; Sun, H.; Hao, L. Transient Stability Assessment of Power Systems Using Cost-sensitive Deep Learning Approach. In Proceedings of the 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 20–22 October 2018; pp. 1–6. [CrossRef]

20. Shinde, B.; Wang, S.; Dehghanian, P.; Babakmehr, M. Real-Time Detection of Critical Generators in Power Systems: A Deep Learning HCP Approach. In Proceedings of the 2020 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, USA, 6–7 February 2020; pp. 1–6.
21. Shao, D.; Xu, Y.; Ma, S.; Jing, Y.; Sun, W.; Zhou, T.; Yang, J. Transient Stability Assessment Method for Power System Based on SVM with Adaptive Parameters Adjustment. In Proceedings of the 2021 IEEE 4th International Electrical and Energy Conference (CIEEC), Wuhan, China, 28–30 May 2021; pp. 1–6.
22. Moulin, L.S.; da Silva, A.; El-Sharkawi, M.; Marks, R.J. Support vector machines for transient stability analysis of large-scale power systems. *IEEE Trans. Power Syst.* **2004**, *19*, 818–825. [CrossRef]
23. 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]
24. Mukherjee, R.; De, A. Development of an Ensemble Decision Tree-Based Power System Dynamic Security State Predictor. *IEEE Syst. J.* **2020**, *14*, 3836–3843. [CrossRef]
25. Baltas, N.G.; Mazidi, P.; Ma, J.; de Asis Fernandez, F.; Rodriguez, P. A Comparative Analysis of Decision Trees, Support Vector Machines and Artificial Neural Networks for On-line Transient Stability Assessment. In Proceedings of the 2018 International Conference on Smart Energy Systems and Technologies (SEST), Seville, Spain, 10–12 September 2018; pp. 1–6.
26. Amraee, T.; Ranjbar, S. Transient Instability Prediction Using Decision Tree Technique. *IEEE Trans. Power Syst.* **2013**, *28*, 3028–3037. [CrossRef]
27. Azman, S.K.; Isbeih, Y.; Moursi, M.; Elbassioni, K. A Unified Online Deep Learning Prediction Model for Small Signal and Transient Stability. *IEEE Trans. Power Syst.* **2020**, *35*, 4585–4598. [CrossRef]
28. Hang, F.; Huang, S.; Chen, Y.; Mei, S. Power system transient stability assessment based on dimension reduction and cost-sensitive ensemble learning. In Proceedings of the 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 26–28 November 2017; pp. 1–6.
29. Li, N.; Li, B.; Gao, L. Transient Stability Assessment of Power System Based on XGBoost and Factorization Machine. *IEEE Access* **2020**, *8*, 28403–28414. [CrossRef]
30. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System; Krishnapuram, B.; Shah, M.; Smola, A.J.; Aggarwal, C.C.; Shen, D.; Rastogi, R., Eds. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.
31. Zhang, N.; Qian, H.; He, Y.; Li, L.; Sun, C. A Data-Driven Method for Power System Transient Instability Mode Identification Based on Knowledge Discovery and XGBoost Algorithm. *IEEE Access* **2021**, *9*, 154172–154182. [CrossRef]
32. Liu, X.; Min, Y.; Chen, L.; Zhang, X.; Feng, C. Data-driven Transient Stability Assessment Based on Kernel Regression and Distance Metric Learning. *J. Mod. Power Syst. Clean Energy* **2021**, *9*, 27–36. [CrossRef]
33. Li, X.; Liu, C.; Guo, P.; Liu, S.; Ning, J. Deep learning-based transient stability assessment framework for large-scale modern power system. *Int. J. Electr. Power Energy Syst.* **2022**, *139*, 108010. [CrossRef]
34. Subrahmanyam, V.S.; Jain, S.; Narayanan, G. Real-time Simulation of IEEE 10-Generator 39-Bus System with Power System Stabilizers on Miniature Full Spectrum Simulator. In Proceedings of the 2019 IEEE International Conference on Sustainable Energy Technologies and Systems (ICSETS), Bhubaneswar, India, 26 February–1 March 2019; pp. 161–166.
35. Zhao, J.; Gómez-Expósito, A.; Netto, M.; Mili, L.; Abur, A.; Terzija, V.; Kamwa, I.; Pal, B.; Singh, A.K.; Qi, J.; et al. Power System Dynamic State Estimation: Motivations, Definitions, Methodologies, and Future Work. *IEEE Trans. Power Syst.* **2019**, *34*, 3188–3198. [CrossRef]
36. Senyuk, M.; Safaraliev, M.; Gulakhmadov, A.; Ahyoev, J. Application of the Conditional Optimization Method for the Synthesis of the Law of Emergency Control of a Synchronous Generator Steam Turbine Operating in a Complex-Closed Configuration Power System. *Mathematics* **2022**, *10*, 3979. [CrossRef]