Influence on Stability Analysis in Distributed Smart Grids Using Computer Aimed Digital Decision Trees

Shen Li[1], Ke Du[2] and Zhouyan Li[3*]

[1,2,3]Safety Supervision Department, State Grid Shanghai Municipal Electric Power Company, Shanghai 200122, Shanghai, China
[1]lishen_1973@163.com,[2]duke_1979@163.com,[3]zhuoyanli2023@163.com

Corresponding Author: Zhouyan Li, zhuoyanli2023@163.com

Abstract. Driven by the multiple forces of digital economy development, environmental protection, and energy structure improvement, both home and abroad have begun to study clean and renewable energy such as distributed power generation as future energy security. Requirements, because it is greatly affected by weather conditions, it is difficult to control and use, often resulting in waste of electric energy, and the impact on the power grid is often unfavorable, but will aggravate the "Matthew effect" of frequency instability. Therefore, we must study how to adapt the power grid to this new situation, and the traditional prevention of the power grid can only be as conservative as possible in the relay protection, but this does not solve its occurrence. Therefore, we tend to use information technology to transmit and communicate to make the power grid intelligent. Through the high-frequency communication of information and the superposition of various other technologies, one plus one is greater than two. This research design chooses to use the decision tree method to realize the stability analysis of the distributed smart grid, to improve energy efficiency, reduce the impact on the environment, improve the security and reliability of power supply, and reduce the power loss of the transmission network. It can also predict the state of the power system in advance, so as to achieve the effect of early detection, early treatment and early resolution.

Keywords:decision tree; distributed; smart grid; stability; power system; Digital Cultural Heritage Influence
DOI: https://doi.org/10.14733/cadaps.2024.S16.178-198

1 INTRODUCTION

Interestingly, our power industry has encountered similar bottlenecks with Europe and America, both in terms of policy to conserve resources and protect the environment to mitigate global warming; the frequent interaction and transactions between supply and demand at the market level, the quiet changes in pricing means are driving the rapid development of smart grid, and will certainly make the future of the power industry a revolution of the sky. Since the definition of smart grid is still unclear, the development direction of China is different from foreign countries, they focus on the direction of distribution grid, because their power system is not well connected; while we focus on the field of transmission grid, because it is in line with the elements of our geographical environment and the relatively close connection of the grid, China's geographical environment is extremely different, vast, and uneven distribution of resources, large-scale load in the eastern Coastal areas, but large-scale clean energy southwest hydropower, three gorges of hydropower, northwest photoelectricity are to be transported through the west-east power transmission project, so special benefit greater ultra-high voltage transmission technology, when they fit perfectly in order to greatly enhance network security and stability. In addition, the way the power grid is developed and the customer's electricity consumption behavior is also changing subliminally at the same time. We should integrate many of the sensors in the transmission grid with phasemeters as the main measurement unit, flexible AC transmission and distributed generation in the supporting supply grid, automated meter reading, demand-side management, and many other technologies[12]. Of course our State Grid has started to implement the SG186 project, which has laid a solid foundation for our intelligent grid. And this huge challenge is requiring us to explore and overcome difficulties to achieve it.Digital cultural heritage can provide historical data and context about the regions and communities where distributed smart grids are implemented. This historical information can be used to supplement the data used in stability analysis, providing a richer understanding of the local environment, potential challenges, and factors that may influence grid stability.

In 2006, IBM proposed the concept of "central neural network" for intelligent power grid planning and primary solutions in conjunction with other large power companies and power industry and related specialized research institutes around the world. The power companies are able to reduce the number of power outages directly through sensors, meters, automatic monitoring devices and other tools, and to manage the power system in a fine-grained manner to restore power as soon as possible. Although proposed early but only in recent years was mentioned again, with the climate conference and a series of issues related to environmental protection has been mentioned, coupled with the need to revitalize the economy the plan has returned to the public eye. Recently, Congress approved a $4.5 billion loan for grid investments, and the smart grid is now seen as a key enabler in helping customers effectively reduce electricity costs and increase the stability of their power supply by collecting, transmitting, aggregating, classifying, and analyzing information using digital technology, and then providing solutions to customers based on the analysis and results of the digital technology, thus This has greatly improved the safety and reliability of the grid and the efficiency of its operations. However, it is clear that local regulators are clearly pessimistic about the investment, and the details will have to be reported in the next data report. In addition, the government has provided more than $2 billion in assistance to U.S. high-tech companies to research star-shaped energy storage battery technology and hybrid vehicles, which are more closely linked to the smart grid; it has also allocated $4 billion to update the U.S. electric infrastructure, which is expected to add 40,000 smart meters and upgrade 4,800 kilometers of transmission lines, as well as develop technologies to reduce line losses.

The Japanese government has invested heavily in supporting and actively promoting the development of the infrastructure and application of an intelligent Japanese power grid, and the main purpose is to gradually form an intelligent Japanese power grid based on its own actual national conditions by focusing on how to significantly invest in the development and utilization of solar energy and how to effectively ensure the stability of the Japanese power grid security operation system. After consultation, a comprehensive test on how to construct an island smart home dedicated grid has been successfully conducted on a small, medium and large scale on an isolated island for all its staff at the same time only in Japan, in order to effectively and uniformly manage the remaining pool power and fluctuations between power supply frequencies on the isolated island

and how to properly use smart batteries under the conditions of large-scale use of solar cell power generation on the isolated island in Japan. The Japanese government has been hoping for the success of the smart power grid application trial and its implementation on a large scale, so that we can effectively boost Japan's domestic demand and create more local jobs by increasing the demand for Japanese power equipment and putting it into use. In order to effectively cooperate with the innovation research of the enterprise's intelligent technology, Tokyo Institute of Technology has planned to set up "comprehensive research institute", in which the research of "smart grid project" about how to combine renewable energy with intelligent power system is also attracting much attention from the industry. The research on how renewable energy can be integrated with smart power systems is also of great interest to the industry. A number of high-tech companies are actively involved in the strategic cooperation with the Tokyo International Power Group, and plan to successfully complete the key technology in at least three and a half years, so that the integrated generation of renewable energy and the existing power system in Japan can be organically integrated to achieve mutual benefits.

Large European national power construction companies carry out intelligent hydroelectric power grid project construction and application practice, due to the current international oil and civil gas market prices are too high, as well as the financial attributes of too strong, and real non-renewable energy, but the contradiction between the large demand for energy and reduce the environmental pressure of greenhouse gas emissions prompted the European countries to build energy inter-regional trade and transmission system, namely "Supergrid". This plan makes full use of the great potential of the great solar wind energy of the Sahara desert in North Africa to complement the future energy system of Europe. In order to better realize this plan, European countries from the beginning of 2009 has been a number of important meetings, the development of supporting policies to support, further clarify the combination of Atlantic wind power, North Africa and southern Europe solar energy better into the European grid. If this is achieved, it will bring about far-reaching changes to the current supply and demand relationship, making it more flexible and cleaner. Europe's research on smart grid technologies is focused on four areas: network assets, grid operations, supply-side and metering, and power generation and customer power reserves; and their goal is to improve operational efficiency, lower electricity prices, and create deeper bonds with customers. The deep collaboration between companies and the grid is also driving technological advances to reduce unnecessary system losses in all four areas of transmission, distribution and use, thereby increasing efficiency. It is expected that in the next ten years or so it will be possible for Europe to reduce emissions by nearly 15%, which will undoubtedly help the EU to complete the 2020 emissions reduction, and also contribute to the development and upgrading of the social economy.

2 RELATED WORK

In the "In 2020-2021 national electricity supply and demand situation analysis" forecast report issued by the China Electricity Council said social electricity consumption grew by an average of 5.7% per year. Grid-connected power generation of the sun and wind power are 261.1 and 466.5 billion kWh respectively, up 16.6% and 15.1% year-on-year respectively. And the distributed smart grid (DSG) electricity is mainly converted from wind and solar energy, i.e. wind energy into electricity and solar energy into electricity. The growth rate of power generation can be seen from the rapid development of distributed power generation.

Nevertheless, large conventional power plants still dominate the grid: transmission lines connect large power plants to regional customers in a local star topology. With the entry of more renewable energy sources, the grid topology becomes more decentralized and cyclical due to distributed generation. In this case, consumers can act as both producers and consumers, and the so-called consumer and electricity transport is no longer one-way. A known challenge for many renewable energy sources is their volatility. Volatility occurs on different time scales, including seasonal fluctuations, day-to-day fluctuations and intra-second fluctuations. This requires fundamental changes in grid control and design strategies, as well as market innovations to ensure cost effectiveness. Therefore, there is a need for more flexible options for supplying and balancing energy in the electricity market, as fluctuating renewable energy sources (RES) cannot guarantee power supply in the same way as conventional power plants. In this regard, it is of utmost importance to identify options that are both cost effective and stabilize the system. So far, the framework for long-term planning of electricity market design and grid stabilization does not meet the need for sufficient flexibility options. Different smart grid approaches have been proposed, suggesting ways to match supply and demand in such a fluctuating grid. However, economic and political feasibility and market integration are often overlooked. A key idea of the various smart grid concepts is to regulate consumer demand, which is a huge paradigm shift compared to the current grid operation schemes [1],[7]. Many proposals for smart grids are based on adequate ICT infrastructure, for example [4],[8]. However, such a centralized system raises issues of cybersecurity and privacy protection, and some studies have highlighted the cost burden associated with these proposals. In contrast, another approach that does not require extensive communication between consumers and producers is to directly use grid frequencies to adjust production and consumption. Frequencies increase when there is excess power and decrease when there is deficit power [10],[11]. Based on 20 earlier ideas, a new smart grid concept, distributed smart grid control (DSGC), was introduced and its mathematical model was presented and analyzed in [8]. Energy consumers using DSGC can control the instantaneous demand based on the grid frequency, which can be easily measured by inexpensive devices. Reference [9] shows that DSGC can improve the stability of grid dynamics if frequency measurements are averaged over a sufficiently long time interval. However, only very small networks have been investigated so far. Therefore, the influence of the grid topology on the grid dynamics under DSGC is a very open research question.

Distributed generation is becoming more and more penetrated in the global grid. It can have a positive impact on the network, but it can also have a negative impact if the integration is not managed properly, with his very obvious problems. This is particularly true for PV, partly because of its highly volatile output and partly because it is being rapidly deployed in many countries. Potential positive impacts on grid operation may include reduced network traffic and thus reduced losses ΔS and voltage drop ΔU. Potential negative impacts of high penetration include voltage fluctuations, voltage rise and reverse currents, power factor changes, frequency regulation, power fluctuations and harmonics, unintentional islanding, fault currents, and grounding problems. And since the frequency fluctuation rate of photovoltaic and wind power is extremely high so it has a great impact on the grid stability and therefore puts a higher demand on the grid control.

DSG and system modeling. Describing the DSG system can be done by using a system of differential equations. There are many limiting factors in studying the stability problem because of the high modeling constraints. We propose to use decision trees (machine learning) to build models and make predictions based on changes in data. This can be done, for example, by using a system of SYS equations to describe the problems that a DSG system will encounter during operation. The fixed input problem due to a single transformation and the problem of homogeneity due to over-idealization.

$$[\tau,\ \ P,\ \ \gamma,\ \ \alpha,\ \ K,\ \ T] \tag{1}$$

Reaction time τ 、Mechanical power P 、Economic elasticity parameters γ 、Motor loss parameters α 、Transmission parameters K and System runtime T Together they formsys Model。

The sys model combines the economic model of function consumption with the "classical model" and "structure maintenance model" in power engineering. Common energy supply nodes and power consumption load nodes in life are mostly composed of asynchronous motors, such as air conditioners, electric fans, etc., and they all obey the law of energy conservation. Therefore, the formula can be approximated：

$$P_s = P_a + P_d + P_t \qquad (2)$$

In the above formula： P_s input power to the motor； P_a is the mechanical power of the rotor； P_d Power loss for the motor； P_t is the system load power.

Putting it into the corresponding physical formula can get：

$$P_{sj} = \frac{1}{2} M_j \frac{d}{dr} (\delta_j)^2 + \mu_j (\delta_j)^2 - \sum_k P_{j,k,\max} \sin(\delta_{k-} \delta_j) \qquad (3)$$

formula：subscript j 、 k Indicates the number of nodes in the system, Same below； r is the vertical distance from any point on the motor (as a rigid body) to the shaft； M is the moment of inertia； μ is the coefficient of friction； δ is the phase angle； $P_{j,k,\max}$ for node j and node k transmission power between.

Configure and simplify the following parameters and bring them into the（4）：

1.transfer parameters $K_{j,k} = \frac{P_{j,k,\max}}{M_j}$ ；

2.Motor Loss Parameters $\alpha_j = \frac{2\mu_j}{M_j}$ ；

3.load parameters $P_j = \frac{P_s - \mu_j \omega^2}{M_j}$, ω is the angular velocity of the rotor;

④ Phase angle parameter $\delta_j(t) = \omega t + \theta_j(t)$, in θ_j is the rotor angle, t for time。

state of each machinej∈{1, …, N}The characteristic is that the rotor is relative to the Ω = 2π × 50 Hz or Ω = 2π × 60 HzThe angle of the rotated grid datum θj（t）and its angular frequency deviation from the referenceωj= dθj/dt. Each machine consists of mechanical powerPi（t）drive, This is positive for generators and negative for consumers. also, Each machine transmits power through adjacent transmission lines with coupling strength Kij. Each machine transmits power through adjacent transmission lines with coupling strength Kij. The dynamics of machine i are given by the equation of motion

$$\frac{d^2\theta_j}{dt^2} = P_j - \alpha_j \frac{d\theta_j}{dt} + \sum_k K_{j,k} \sin(\theta_k - \theta_j) \qquad (4)$$

where α is the damping constant. We ignore the ohmic load, which should be 33% smaller than the shunt admittance for the dynamics we consider. We consider the moment of inertia to be the same for all machines and therefore remove the moment of inertia from the equations of motion for ease of representation.distributed Smart Grid Control is based on demand response and aims to stabilize the electricity system by encouraging consumers to reduce consumption during times of high load and low production, and to increase consumption during times of low load and high production. Instead of paying a constant price for electricity. Consumers present a linear price-frequency relationship. Among them, configure the elastic scale factor c_p

$$p_{rj} = p_{r\omega} - c_p \int_{t-T_j}^{t} \frac{d\theta_j}{d_t}(t-\tau_j)dt \tag{5}$$

$$\tilde{P}_j(p_i) = p_j - c_j(p_{rj} + p_{r\omega}) \tag{6}$$

In the formula： p_{rj} for node j electricity price； $p_{r\omega}$ for $d\theta_j / dt \equiv 0$ electricity price； $\tilde{P}_j$ is the corresponding output of the node at the moment corresponding to the electricity price. c_j is the price elasticity ratio coefficient； τ_j is the reaction time of the node； T_j is the average time of the node.

式（2.4）—（2.6）The formula obtained by combining and simplifying the calculation：

$$\frac{d^2\theta_j}{dt^2} = P_j - \alpha_j \frac{d\theta_j}{dt} + \sum_{k=1}^{N} K_{j,k} \sin(\theta_k - \theta_j) - \frac{c_p c_j}{T_j}(\theta_j(t-\tau_j) - \theta_j(t-\tau_j - T_j)) \tag{7}$$

Defining Economic Elasticity Parameters $\gamma_j = c_p c_j$, get

$$\frac{d^2\theta_j}{dt^2} = P_j - \alpha_j \frac{d\theta_j}{dt} + \sum_{k=1}^{N} K_{j,k} \sin(\theta_k - \theta_j) - \frac{\gamma_j}{T_j}(\theta_j(t-\tau_j) - \theta_j(t-\tau_j - T_j)) \tag{8}$$

After rationally configuring the physical parameters of different nodes of each system according to the established design scheme, Get the parameter equations for each node of the dsg system.

After the distributed modeling process described earlier, we need to find out an index to determine its local linear stability, It is the value of the real part of the root of the system characteristic equation S_R , This indicator can determine whether the system is unstable：when S_R >0, System instability requires manual intervention；when S_R <0, The system remains stable and everything works fine; when S_R =0, It is a critical state and requires personnel to go to check. So find out S_R The value of is critical to us.

Star DSG modeling. The relevant data in the operation of the system uses the parameters of the equation data set shown in Table 2.1 from UCI. By changing the different input states of the equation, the relevant data and operation stability conclusions are obtained. The system operation stability and data simulation model Such as a star power supply system between 4 nodes.

variable	describe	type
τ	elastic reaction time	enter
P	Node Power Parameters	enter
γ	Economic Elasticity Parameters	enter
α	Motor Loss Parameters	Steady
K	transfer coefficient	Steady
T	Pricing cycle factor	Steady
S_R	system stability criterion	output

Table 1: DSGSystem Parameter Description.

A total of 10,000 sets of output data are obtained after judging the value of the owned value by the criterion. These 10,000 sets of data will be analyzed and trained to train the model. After statistics, there are a total of 6380 groups of unstable samples, and a total of 3620 groups of stable samples.

sample	τ_1	τ_2	τ_3	τ_4	P_1	P_2	P_3	P_4	γ_1	γ_2	γ_3	γ_4	S_R
1	2.96	3.07	8.38	9.78	3.76	-0.78	-1.25	-1.72	0.65	0.85	0.88	0.95	0.055
2	9.30	4.90	3.04	1.36	5.06	-1.94	-1.87	-1.25	0.41	0.86	0.56	0.87	-0005
3	8.97	8.84	3.05	1.21	3.40	-1.20	-1.28	-0.92	0.16	0.77	0.83	0.11	0.003
...							...						
10000	6.53	6.78	4.34	8.67	3.49	-1.39	-1.53	-0.57	0.07	0.51	0.38	0.94	0.045

Table 2: DSG Stability Dataset.

In the follow-up, the CART method for decision trees will be used to conduct research and data analysis on this issue.

3 DECISION TREES FOR MACHINE LEARNING

3.1 A Brief Introduction to Machine Learning

Machine learning technology is an interdisciplinary subject, which covers the basic theoretical knowledge of multiple disciplines such as probability, statistics, algorithms and complexity. It originated from the fact that more and more data sets are used for human measurement, and it is still not enough to solve the problem under the condition of initial manpower plus tools. Therefore, in order to solve such engineering problems, scientists began to study machine learning and wanted to use computers to replace manual work. Conduct predictive learning to solve problems. The subject of machine learning mainly focuses on the study of how electronic computers learn and behave human beings through programming to acquire new knowledge, and there is a continuous self-learning and pre-updating ability to approach the ideal goal, so it is the core of artificial intelligence.The advantage is that it saves a lot of manpower, and once the model structure is completed, the prediction speed is extremely fast and the accuracy is high. And it has a wide range of applications. The disadvantage is that still relying on manual construction of feature engineering, there will be omissions in feature selection, or addition of invalid features.

3.2 Fundamentals of Decision Trees

Decision tree is a common algorithm in machine learning. It can be regarded as one of the basic methods for beginners. It originated from the ID3 algorithm and C4.5 algorithm proposed 20 or 30 years ago, and has been continuously developed. It can be roughly subdivided into various judgmental variable classification decision trees and variable regression decision trees. The former's various judgmental variable classification decision criteria are usually discrete and independent, and are not only applicable to regression decision trees. Some parameters entered in one variable are more suitable for continuous variables; the two decision trees have a series of advantages such as easy to read and understand, obvious and fast classification, and easy to use.

Decision trees are graphs or tables that help determine a course of action or show statistical probabilities. The diagram is called a decision tree because of its resemblance to the plant of the same name, usually outlined as an upright or forked horizontal figure. Starting with the decision itself (called a "node"), each "branch" of a decision tree represents a possible decision, outcome, or response. The furthest branches on the tree represent the end result of a decision path and are called "leaves". People use decision trees to clarify, map and find answers to complex questions. Decision trees are often used to determine a course of action in finance, investing, or business. In mathematics, decision trees are also known as dendrograms.

3.3 Decision Tree Modeling and Learning

The essence of decision tree learning is to predict the value of the target variable by creating a classification model trained by several input variables. A decision tree is a simple representation of a classification example, usually an approximation of the optimal solution to a problem. Its learning algorithm includes the identification and selection of features (selecting the best features from many features), the generation and pruning of decision trees. Commonly used algorithms for decision trees are ID3, C4.5, CART[5],[3],[2],[13],[14].

Information entropy H（x）: In the 1940s, information entropy was proposed by Shannon to quantify the quality of an information. The basic definition of entropy is:

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i \tag{9}$$

Among them, x represents a finite discrete function random variable used to describe a finite number of random values, and pi represents the change probability of each discrete random variable during the entire random event occurrence time. The mean value of information entropy before classification without other loan data is:

$$H(X) = -\frac{6}{15}\log\frac{6}{15} - \frac{9}{15}\log\frac{9}{15} = 0.971 \tag{10}$$

The magnitude of entropy indicates the uncertainty of a random variable. The greater the entropy, the greater the unknown information, and the smaller the entropy, the smaller the unknown information. If everyone needs to apply for this information entropy, the result is an information entropy as follows:

$$H(X) = -\frac{15}{15}\log\frac{15}{15} = 0 \tag{11}$$

That is, the result is known. One of its ultimate goals is to minimize the information entropy and judge an event with the highest probability through these characteristics.

Conditional entropy H（Y|X）：Represents the uncertainty of random variable Y under the condition of known random variable X, defined as:

$$H(Y \mid X) = \sum_{i=1}^{n} p_i H(Y \mid X = x_i) \tag{12}$$

pi=P（X=xi），That is, the probability of xi in the variable；H（Y|X=xi） X=xi is Yentropy. as shown in the picture, X is age,

in the same way

$$H(Y \mid X = 2) = 0.971 \tag{13}$$

$$H(Y \mid X = 3) = 0.722 \tag{14}$$

The final conditional entropy is：

$$H(Y \mid X) = \frac{1}{3}*0.971 + \frac{1}{3}*0.971 + \frac{1}{3}*0.722 = 0.888 \tag{15}$$

information gain g（y，x）:is used to indicate that for a known type y The information of other types of y leads to a degree of reduction of information uncertainty of other types of y, and the definition formula is:

$$g(Y, X) = H(Y) - H(Y \mid X) \tag{16}$$

Among them, H(Y) is the empirical entropy of sample category Y, and H(Y|X) is the empirical conditional entropy. The above figure is an example:

$$g(Y,X)=H(Y)-H(Y\mid X)=0.971-0.888=0.083 \tag{17}$$

This method is mainly used in the ID3 algorithm, but there is a big loophole in ID3, that is, the selection of multiple features is ignored, which will have a great impact on the results. In extreme cases, a huge and There are deep and shallow trees, which will cause very unreasonable consequences to the results, so c4.5 has been improved, using the information gain ratio gR (Y, X)

$$g(Y,X)=H(Y)-H(Y\mid X) \tag{18}$$

Gini index: In a factor classification calculation problem, assuming that a sample point contains at least k classes containing the number of factors, and the probability distribution of each sample point at kth is pk, then the Gini of the distribution of this probabilistic factor class The index can be defined by the formula:

$$Gini\ (p)=\sum_{k=1}^{K} p_k(1-p_k)=1-\sum_{k=1}^{K} p_k^2 \tag{19}$$

The Gini index is similar to entropy, both of which are used to represent the uncertainty of the sample. It is also a basis for judgment. The Gini index is similar to the entropy, both of which are used to represent the uncertainty of the sample, but the Gini coefficient is mainly used in classification and regression trees.

4 REGRESSION TREE（CART）

4.1 CART

CART, classification regression tree, is the basis of almost all complex decision tree algorithms, and has the following characteristics:

1.CART is a binary tree

2.CART can be both a classification tree and a regression tree, and the specific situation depends on the target task

3.When the value of cart is a regression classification node tree, the value of GINI can be used as the main node of the classification node tree and the calculation basis for regression splitting; for example, when the value of cart is a node regression splitting tree When , the addition of MSE (mean square error) can be used as the main calculation basis for classification node regression splitting.

4.2 Difference Between Classification Tree and Regression Tree

For classification tasks, it is a classification tree; for regression tasks, it is a regression tree. Classification task: The prediction target is usually a discrete value. For example, we want to predict whether the user may be overdue in the process. If it exceeds this, it is one kind, and it is represented by 1. If it does not exceed this, it is another kind, and it is 0. express. The classification tree can use the GINI value as the basis and basis for splitting nodes;

Regression target task: The regression target task values that can be predicted are usually continuous, such as predicting a user's height [7],[4]. Regression subtree splitting can be performed by using foreign functions such as mse (mean square error) as the number of regression nodes.

4.3 Detailed Algorithm

Let sample S set : Samples={s1 ,s2,..., s i}, is integrated by having i samples.

Let feature F set : Features={f1,f2,...,f j}Each sample corresponds to a set of features.

target value/true valuegather : T ={t 1,t 2,...,t N }, Each sample corresponds to a target value. For regression tasks, each target value is a value with a continuous value attribute.

Proceed as follows:

1. The original data set X, at this time the depth of the tree depth=0.

2.For the set X, traverse each value of each feature, and use the value to split the original data set S into two sets: left set Sleft (<=value sample), right set Sright (>value sample), each A collection is also called a node. Calculate the mse of these two sets separately, find the value that makes (leftmse+rightmse) the smallest, record the feature (feature) name and value at this time, this can be expressed as the best segmentation feature and the optimal segmentation of this feature value;

The calculation method of the mean square error (mse) of each set/node N points is as follows:

1.

$$\text{mean} = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{20}$$

where N is the total number of samples in the set, y i is the target value of each sample in the set (special note: this function mean is a predicted value for the predicted node, that is to say, two samples scattered in the predicted node may have the same result Two samples within the same predicted node may have the same node predicted value.)

2.

$$\text{mse} = \sum_{i=1}^{N} (y_i - mean)^2 \tag{21}$$

As long as it is a mathematical formula that can accurately measure the correlation gap between the predicted value and the real target value/site price, it can be used, such as information gain, information gain ratio, Gini coefficient, etc. But the mean square error has better benefits: the first-order derivative and the second-order derivative can be found and easy to find. This is a very important point.

3. After finding the best segmentation feature and the best segmentation value, use the value to split the set S into two sub-sets: the left set Sleft and the right set Sright, and each set is also called a node. The depth of the tree at this timedepth += 1.

4. For collection S right S leftRepeat steps separately2, 3, until the termination condition is met.

5 SIMULATION AND EXPERIMENTAL RESULTS ANALYSIS

5.1 Program Simulation Writing Process

Open the csv file, and you can see that there are 12 columns of data (as shown in Figure 4.1). The initial parameters are tau1, tau2, tau3, tau4, p1, p2, p3, p4, g1, g2, g3, g4. The actual meanings of these distributions are the initial parameters of the nodes are active power P, reactive power Q, voltage U, power angle δ, and tau, p, g represent three distributed nodes. Obviously, tau1 represents the active power detection parameter of the tau node, and g3 represents the monitoring parameter of the voltage of the g node. The m column is the value corresponding to the prediction result, and the n column represents whether the prediction result is stable or unstable at this time. From the previous criterion, we can see that when S>0, it is in an unstable state, which is 0; When S<0, it is in stable state, which is Fig. 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	tau1	tau2	tau3	tau4	p1	p2	p3	p4	g1	g2	g3	g4	stab	stabf	stabf1
2	2.95906	3.079885	8.381025	9.780754	3.763085	-0.7826	-1.25739	-1.72309	0.650456	0.859578	0.887445	0.958034	0.055347	unstable	0
3	9.304097	4.902524	3.047541	1.369357	5.067812	-1.94006	-1.87274	-1.25501	0.413441	0.862414	0.562139	0.78176	-0.00596	stable	1
4	8.971707	8.848428	3.046479	1.214518	3.405158	-1.20746	-1.27721	-0.92049	0.163041	0.766689	0.839444	0.109853	0.003471	unstable	0
5	0.716415	7.6696	4.486641	2.340563	3.963791	-1.02747	-1.93894	-0.99737	0.446209	0.976744	0.929381	0.362718	0.028871	unstable	0
6	3.134112	7.608772	4.943759	9.857573	3.525811	-1.12553	-1.84597	-0.55431	0.79711	0.45545	0.656947	0.820923	0.04986	unstable	0
7	6.999209	9.109247	3.784066	4.267788	4.429669	-1.85714	-0.6704	-1.90213	0.261793	0.07793	0.542884	0.469931	-0.01738	stable	1
8	6.710166	3.765204	6.929314	8.818562	2.397419	-0.61459	-1.20883	-0.574	0.17789	0.397977	0.402046	0.37663	0.005954	unstable	0
9	6.953512	1.379125	5.7194	7.870307	3.224495	-0.749	-1.18652	-1.28898	0.371385	0.633204	0.732741	0.380544	0.016634	unstable	0

Figure 1: Raw data table.

5.2 Experimental Analysis Results

Because the data set is too large, we take 10 sets of samples to judge, and the data of and is from 60 lines to 70 lines. 1. When depth=4, the original data in these ten sets of data are -0.009, 0.072, -0.021, 0.060, 0.062, 0.071, -0.009, 0.044, 0.011, -0.048. The corresponding original state is 4 stable and 6 unstable; the predicted data obtained after program training are 0.001, -0.059, 0.044, 0.045, -0.042, -0.030, -0.027, 0.007, -0.044, 0.011. Five of the predictions were stable and five were unstable. At this point the overall prediction of the model is accurate: 29.222%, very inaccurate. The reason for the analysis is that the number of node branches is only four layers, and the amount of data is too large. It can be seen from the figure below (4-3) that the data distribution is uneven, with as few as sixty or seventy samples and as many as hundreds or thousands directly lead to inaccurate results. Therefore, in the figure below, the number of node layers should be increased to improve the prediction accuracy.

Raw data	*forecast data*	*original result*	*forecast result*
-0.009	*0.001*	*stable*	*unstable*
0.072	*-0.059*	*unstable*	*stable*
-0.021	*0.044*	*stable*	*unstable*
0.060	*0.045*	*unstable*	*unstable*
0.062	*-0.042*	*unstable*	*stable*
0.071	*-0.030*	*unstable*	*stable*

-0.009	*-0.027*	*stable*	*stable*
0.044	*0.007*	*unstable*	*unstable*
0.011	*-0.044*	*unstable*	*stable*
-0.048	*0.011*	*stable*	*unstable*

Table 3: Depth=4d; corresponding prediction results.

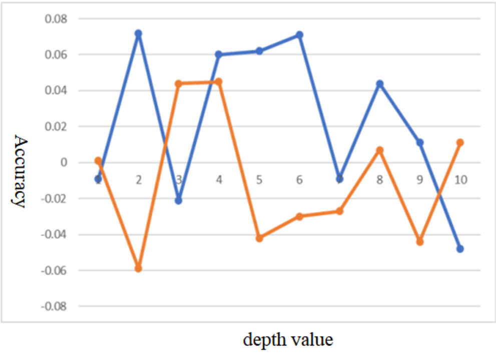

Figure 2: Depth=4; corresponding line table.

The decision tree model formed at this time is shown in the figure:

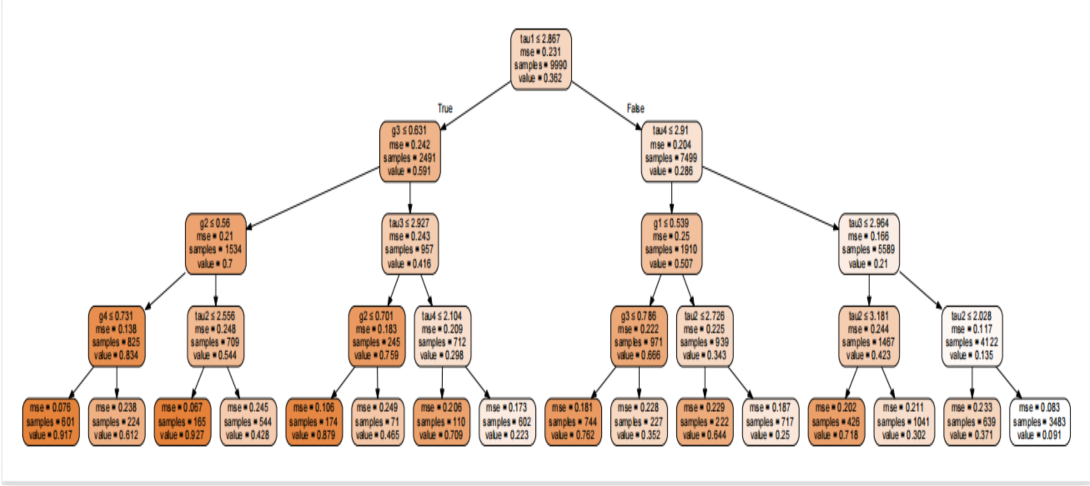

Figure 3: Depth=4; tree.

The original data in these ten sets of data are the same as the original results; the predicted data obtained after program training are 0.014, -0.011, 0.071, 0.009, -0.044, -0.011, -0.048, -0.027, 0.026, 0.042. Four of the predictions were stable and six were unstable. At this time, the model's overall prediction is accurate: 39.555%, which is also very inaccurate. The reason for the analysis is that the number of node branches becomes 5 layers, and the result is inaccurate. Therefore, it is still necessary to increase the number of node layers in the figure below to improve the prediction accuracy.

Raw data	*forecast data*	*original result*	*forecast result*
-0.009	*0.014*	*stable*	*Unstable*
0.072	*-0.011*	*unstable*	*Stable*
-0.021	*0.071*	*stable*	*unstable*
0.060	*0.009*	*unstable*	*unstable*
0.062	*-0.044*	*unstable*	*stable*
0.071	*0.011*	*unstable*	*unstable*
-0.009	*-0.048*	*stable*	*stable*
0.044	*-0.027*	*unstable*	*stable*
0.011	*0.026*	*unstable*	*unstable*
-0.048	*0.042*	*stable*	*unstable*

Table 4: Depth=5; data and corresponding prediction results.

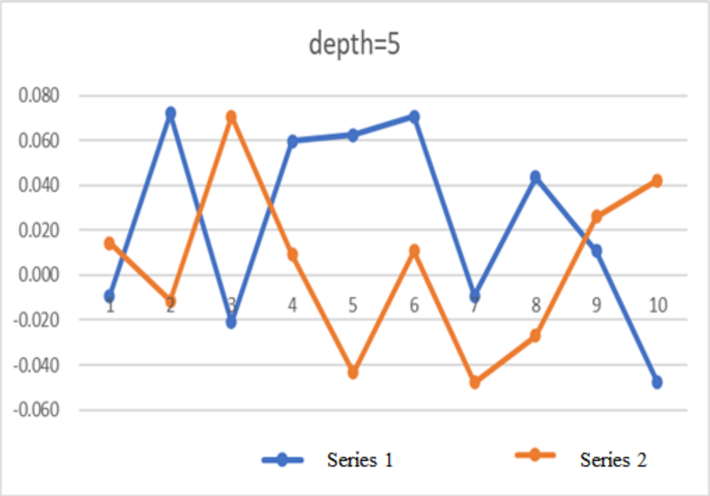

Figure 4: Depth=5; corresponding broken line table.

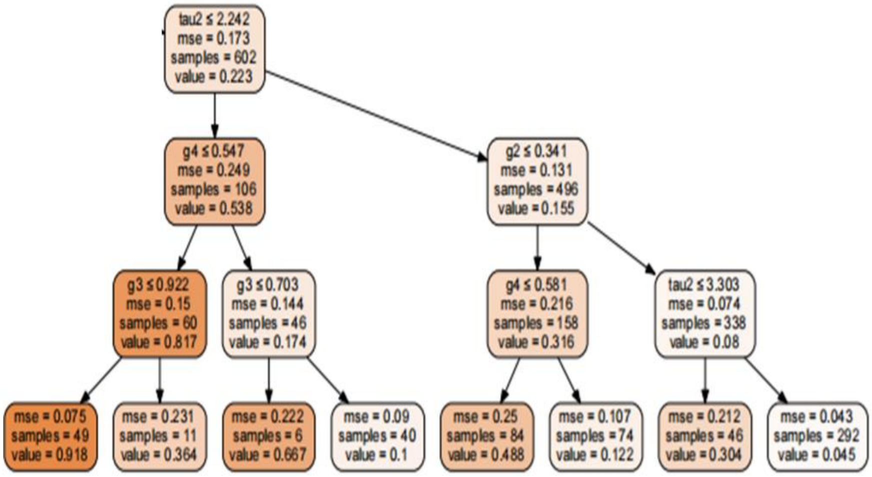

Figure 5: Depth=5; time dendrogram.

When depth=6, the original data in these ten sets of data are the same as the original results; the predicted data obtained after program training are 0.029, 0.050, 0.015, 0.006, 0.017, 0.073, 0.012, 0.073, 0.028, 0.053. Ten of the predicted results were unstable. At this time, the overall prediction of the model is accurate: 64.318%, which is a huge improvement compared to the previous one.

Raw data	*forecast data*	*original result*	*forecast result*
-0.009	*0.029*	*stable*	*unstable*
0.072	*0.050*	*unstable*	*unstable*
-0.021	*0.015*	*stable*	*unstable*
0.060	*0.006*	*unstable*	*unstable*
0.062	*0.017*	*unstable*	*unstable*
0.071	*0.073*	*unstable*	*unstable*
-0.009	*0.012*	*stable*	*unstable*
0.044	*0.073*	*unstable*	*unstable*
0.011	*0.028*	*unstable*	*unstable*
-0.048	*0.053*	*stable*	*unstable*

Table 5: Depth=6; corresponding prediction results.

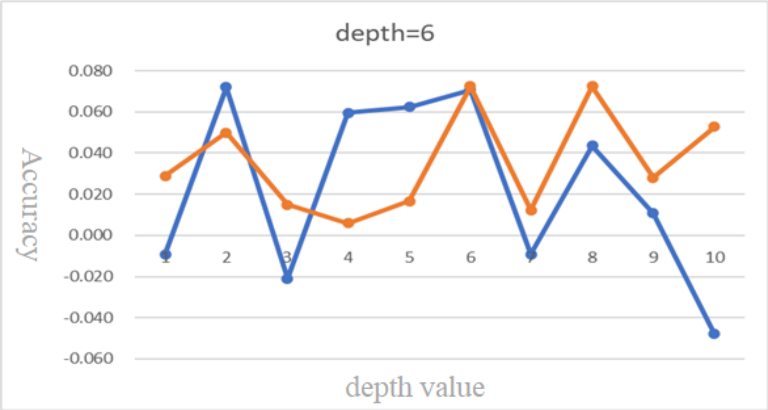

Figure 6: Depth=6; corresponding broken line table.

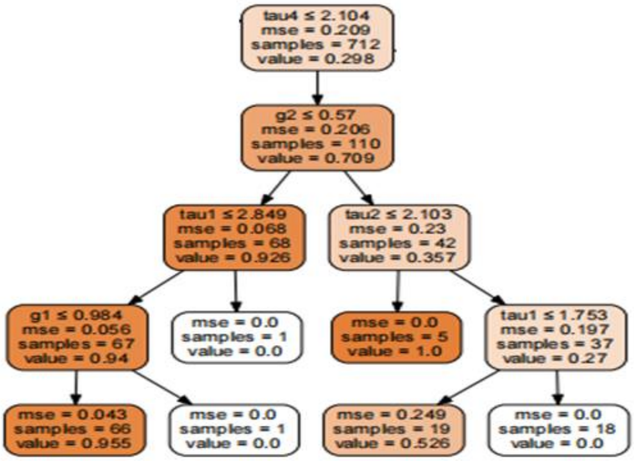

Figure 7: Depth=6; tree.

When depth=7, the original data in these ten sets of data are the same as the original results; the predicted data obtained after program training are 0.042, 0.011, -0.003, 0.028, 0.031, 0.012, -0.018, 0.040, 0.078, 0.035. Two of the predictions were stable and eight were unstable. At this time, the overall prediction of the model is accurate: 87.367%, which is a huge improvement compared to the previous one. The reason is that the number of node branches becomes 7 layers, and the accuracy of the prediction results is already good.

Raw data	*forecast data*	*original result*	*forecast result*
-0.009	*0.042*	*stable*	*unstable*
0.072	*0.011*	*unstable*	*unstable*

-0.021	*-0.003*	*stable*	*stable*
0.060	*0.028*	*unstable*	*unstable*
0.062	*0.031*	*unstable*	*unstable*
0.071	*0.012*	*unstable*	*unstable*
-0.009	*-0.018*	*stable*	*stable*
0.044	*0.040*	*unstable*	*unstable*
0.011	*0.078*	*unstable*	*unstable*
-0.048	*0.035*	*stable*	*unstable*

Table 6: Depth=7; data and corresponding prediction results.

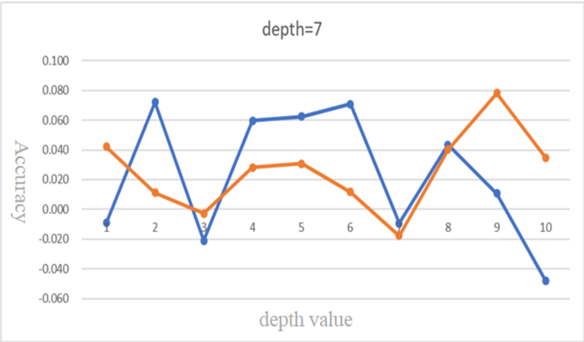

Figure 8: Depth=7; corresponding polyline.

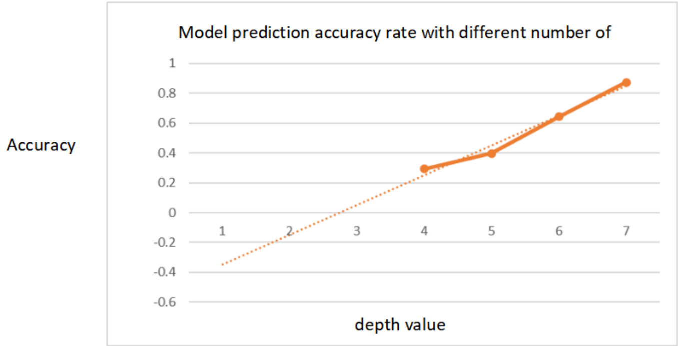

Figure 9: The prediction accuracy of the model.

4.10 The prediction accuracy of the model when the number of nodes is different.It is not difficult to see that as the number of depth nodes increases, the accuracy rate increases. When depth=4, the prediction accuracy rate reaches 29.222%%. When depth=7, the prediction accuracy rate reaches 87.367%. It shows that the increase of branch nodes makes the model more complex and increases the prediction accuracy of the model, but it does not mean that the more branches the node has, the more specific cases should be considered. The data used in this article is only 10,000 groups, and each group has 15 data. The size of this data packet is considered medium, so the time consumed by the machine program operation is not particularly long, but in fact it is not necessary, and when the data packet is particularly large is suitable for multiple nodes.

5.3 Identification and Analysis of Decision Tree Diagram

To analyze how to identify the generated decision tree graph, we take the graph when depth=7 as an example, that is, graph (4.9). Since it needs to be split 7 times, a maximum of 27=128 branches will be generated, but it is divided into the following cases:

1. The feature has been used up and the split cannot continue;
2. The samples of the separated nodes are already of the same class, so there is no need to distinguish them;
3. When the number of samples is 0, it cannot be split.
4. depth >= max_depth, tree stops splitting.
5. The sample size of the node<min_samples_leaftree stops splitting.

Of course, because the obtained result graph is too narrow and long, we can only divide it into several segments and set them as A, B, C, D, E... segments.

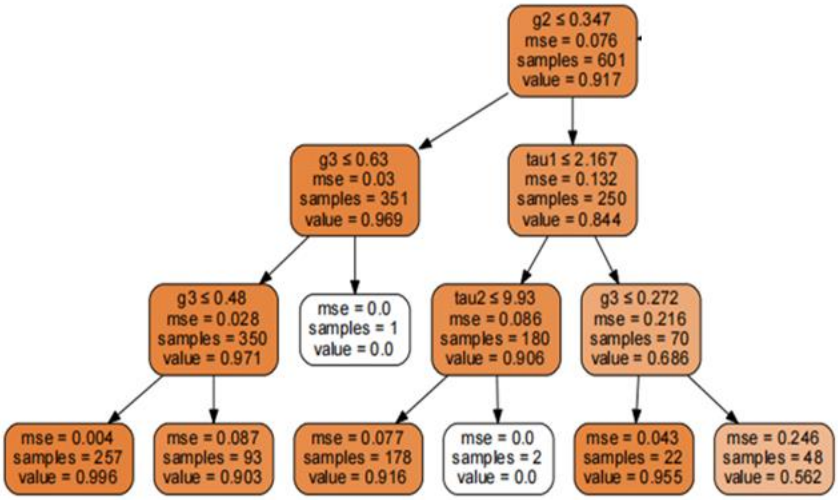

Figure 10: Decision tree section A.

The uppermost criterion for separating nodes in the figure above is that g2 is less than or equal to 0.347. From the explanations in previous chapters, we know that g2 represents the reactive power Q of node g, mse represents the set of N variances at this time, and samples represent samples The number of the value is 601. After the value finds the best segmentation feature and the best segmentation value, the value is used to split the set S into two sets, namely S left and S right. Why choose g2, because here we divide the 12 features into different groups of feature cut points and then calculate the variance sets corresponding to different cut points at this time, compare each other and take the minimum value, which is the value of mse at our node, which is exactly equal to 0.076 , which is indeed a very, very small value numerically. The set of N variances we get under this feature is only the smallest, and the corresponding value at this time is exactly 0.917.

Next, we repeat the previous work to calculate the variance of each node and calculate the mse of the next few places. It is obvious that 601 is divided into two parts. The feature on the side is that tau1 is less than or equal to 2.167, mse=0.132, the number of samples is 250, and value=0.844. Next, we will not repeat the right branch. In the next layer, the feature on the left is still g3, but the value is different. This time g3 is less than or equal to 0.48, mse=0.028. Since the feature is close to the upper layer, the number of samples here is 350, which is very close, value=0.971 ;The right branch is white, indicating that there is 1 sample that cannot be branched here, and it is obvious that the number of samples of the node here has reached the threshold: if the number of samples of a set (node) is less than the minimum threshold number of samples on the left branch, then The tree stops splitting. The disappearance of the features in the seventh layer indicates that they all share one feature at this time—the voltage of the g node, the variance sum on the left is 0.04, a total of 257 samples, the corresponding value is 0.996, and the variance sum on the right is 0.087, a total of 93 samples, the value is 0.903. Calculating from left to right, we can calculate all the parameters, the minimum variance sum at N points (the most suitable feature segmentation point), the corresponding number of samples and the best segmentation value.

This deduction does not calculate the subsequent part of the decision tree. Since the amount of manual calculation is very large (should be the calculation part), the subsequent deduction is all the same as above.

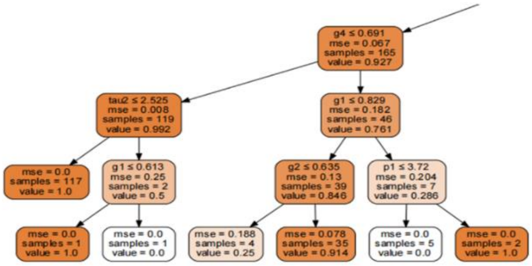

Figure 11: Decision tree section C.

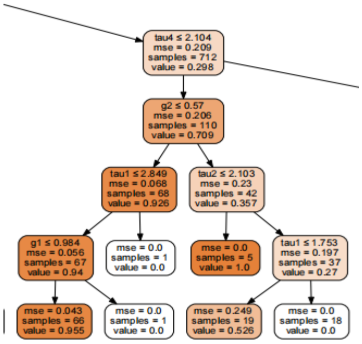

Figure 12: Decision tree section H.

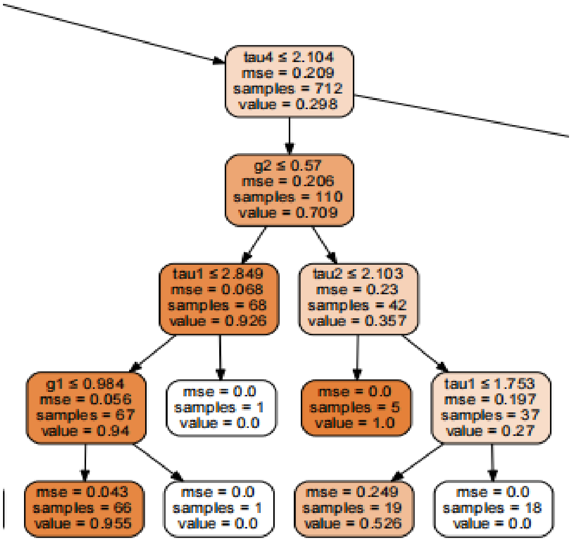

Figure 13: Decision tree section H.

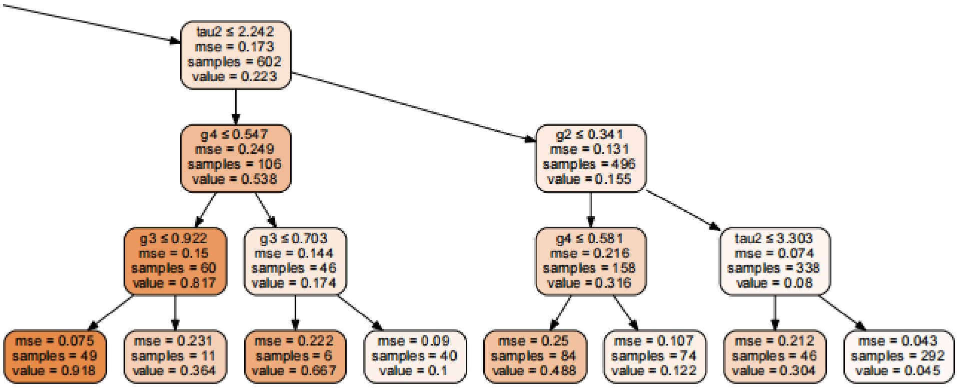

Figure 14: Decision tree section I.

6 CONCLUSIONS

It mainly introduces the status quo and development of distributed (new energy) power generation and smart grid, and introduces the development status, differences and improvement of my country's development of new energy to keep up with international development under the high pressure of environmental protection and international geopolitics my country's energy structure. It also introduces the policies, measures, and implementation effects of various countries, so that my country can extract the essence and discard the dross, so as to lay a solid foundation for the future transformation. This paper mainly studies the influence of distributed generation on the stability of smart grid, which puts forward higher requirements for the future smart grid. Discussed the modeling process of distributed power generation nodes, judged stability indicators (criteria), and knew the knowledge of various data parameters; and explained the difficulties of traditional stability judgment methods, and through the use of differential equations There will be many restrictive factors in the study of stability problems in the way of groups, and there are many modeling constraints and the engineering volume is huge and complex. Use the method of decision tree to predict and judge the stability of the system. Boldly and innovatively combining machine learning with power grid stability analysis, from another perspective, the pre-judgment of the stability of the smart grid can be pre-processed in advance to reduce the processing cost when a certain node fails. There are still some deficiencies in this design, mainly in that the experimental model is relatively simple, the interference factors are relatively rare, and the data processing form is too simple and so on.

Shen Li, https://orcid.org/0009-0007-9143-7433
Ke Du, https://orcid.org/0009-0002-9068-9632
Zhouyan Li, https://orcid.org/0009-0009-5095-5714

REFERENCES

[1] Albadi, M. H.; El-Saa, Da Ny E. F.: A Summary of Demand Response in Electricity Markets, Electric Power Systems Research, 78(11), 2008, 1989-1996. https://doi.org/10.1016/j.epsr.2008.04.002
[2] Chaoyang, Q.; Yufeng, G.: Xin, N.: Design of Network Fault Diagnosis Expert System Model Based on Decision Tree, Jilin Electrical Engineering Society.
[3] Chengzhi, J.; Mingfeng, Y.; Zeyuan, G.: Intelligent Robot Control Method Based on Decision Tree Learning, Computer Engineering and Design, 24(10), 2002, 24-26.
[4] Kok, J. K.; Warmer, C, J.; Kamphuis, I. G.: PowerMatcher: Multiagent Control in the Electricity Infrastructure, International Joint Conference on Autonomous Agents & Multiagent Systems. ACM, 2005. https://doi.org/10.1145/1082473.1082807
[5] Lin, L.; Guangping, T.; Feng, Y.: Application Research of Data Mining Method Based on Decision Tree in CRM, Computing Technology and Automation, 025(001), 2006, 67-69.
[6] Luning, J.: Occam's Razor and Information Security Management, China Information Security, 20(7), 2016, 91
[7] Palensky, P.; Dietrich, D.: Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads, IEEE Transactions on Industrial Informatics, 7(3), 2011, 381-388. https://doi.org/10.1109/TII.2011.2158841
[8] Rosinger, C.; Uslar, M.: Smart Nord Final Report.
[9] SchFer, B.; Matthiae, M.; Timme, M.: et al. Decentral Smart Grid Control, New Journal of Physics, 17(1), 2015. https://doi.org/10.1088/1367-2630/17/1/015002
[10] Schweppe, F. C.: Frequency adaptive, Power-Energy Re-Scheduler: US, US4317049 A, 1982.
[11] Short, J. A.; Infield, D. G.; Freris, L. L.: Stabilization of Grid Frequency Through Dynamic Demand Control, IEEE Transactions on Power Systems, 22(3), 2007, 1284-1293. https://doi.org/10.1109/TPWRS.2007.901489
[12] Xu Han, yo K.: Looking at the Benefits of Electricity Use in the Non-Ferrous Industry from the Perspective of Power Transformation, China Nonferrous Metals, 24(08), 2019, 44-45.
[13] Yilong, G.: Research on Voltage Stability Evaluation of Power System Based on Decision Tree, Daqing:Northeast Petroleum University,2018.
[14] Zhixian, P.: Research on Power Load Forecasting Model Based on Decision Tree, Beijing:North China Electric Power University, 2011.