Algorithm analysis based on machine learning in alarm information of metering system

Yuhan Ma*
School of Management, Tianjin University of Technology, Tianjin, China
*Corresponding author e-mail: yuhanma@tjut.edu.cn

Abstract. Due to the huge scale of the metering warning system and the wide variety of warning information, manual identification can no longer meet the real-time monitoring needs of modern power grids. Through the use of machine learning, data mining technology, linear regression method, naive Bayes method, automatic classification analysis of alarm information from the massive alarm information. Construct an SVM prediction model to improve the accuracy of alarm information analysis and improve the effectiveness of power grid monitoring and management. The experimental results show that it can provide support for the normal operation of the grid metering system.

1. Introduction
With the rapid development of computer technology, network technology, and database technology, power grid operation management has developed many automation systems, such as SCADA system, EMS system, OPEN-3000 management system, remote meter reading system, etc. This can collect information from the power grid. Data is an integral part of the grid operation management measurement system [1]. At present, the metering system, as one of the critical contents of the power grid information system, can maintain the correct operation of the power system, quantify the power information, and realize the automatic dispatch and management of power equipment and the network [2]. However, due to the rapid development and construction of the State Grid's operating network, the network topology is complex, and the types and quantities of access equipment are tens of thousands. Therefore, building a professional on-duty team is necessary to conduct comprehensive remote monitoring and analyze monitoring information around the clock and across the region. The workload is enormous, which brings severe work pressure to the staff, and reduces the quality and accuracy of the analysis of monitoring information. In order to be able to process the alarm information of the metering system quickly, the alarm monitoring signal can be logical, organized and effectively analyzed. An intelligent analysis tool can be provided for the monitoring personnel to strongly support the judgment and decision-making of the staff on duty and locate the measurement system. The fault location improves the reliability and accuracy of grid operation [3].

Machine learning algorithms can mine and classify and manage the massive alarm information of the metering system to help monitors discover valuable information, realize comprehensive real-time information collection and analysis and processing on the power demand side, and business analysis and processing based on electricity metering. The research can support the construction of the distribution network, improve the power utilization efficiency of the power consumption terminal, and enhance the
comprehensive social benefits. This paper analyzes the application and development status of alarm information in a metering system in detail, describes commonly used machine learning algorithms, and designs an alarm information mining algorithm based on SVM. Experimental results show that the algorithm can accurately mine metering information data patterns, improve the accuracy of power grid monitoring, and optimize data processing content in a shorter data processing time.

2. Analysis of commonly used machine learning algorithms

2.1. Linear regression algorithm

In the alarm information analysis process of the measurement system, the linear regression algorithm adopts the objective function and the error function to realize the classification of data information. The objective function is shown in formula (1) and formula (2).

\[
y(x,w) = w_0 + \sum_{j=1}^{m-1} w_j \Phi_j(x)
\]

\[
E(w) = E_D(w) + E_W(w) = \frac{1}{2} \sum_{n=1}^{N} \{t_n - w^T \Phi(x_n)\} + \frac{\lambda}{2} w^T w
\]

Among them, \( \Phi_j(x) \) represents a basis function. The function can choose the Gaussian function or polynomial function. Linear-Gaussian model and polynomial fitting model can be used in the classification process. \( E_D(w) \) represents the error generated by predicting the target value, \( E_W(W) \) represents the penalty function of the parameter, \( \lambda \) is called the regular coefficient, \( w_0 \) represents the deviation parameter, and the incidence matrix \( \Phi \) is a column whose values are all 1. Linear regression algorithm can use the maximum likelihood method to find the solution of \( w \), the form of the solution of \( w \) is shown in formula (3).

\[
W_{ML} = (\Phi^T \Phi)^{-1} \Phi^T t
\]

2.2. Naive Bayes Method

Linear regression algorithm data classification solution is complicated, which is not conducive to implementation. Therefore, a supervised learning algorithm can train the sample set to generate a data analysis model, such as the naive Bayes algorithm. The specific objective function is formalized as follows:

Suppose the training sample set \( D=\{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\} \) contains N samples, where \( x_i=(x^{(1)}_i,x^{(2)}_i,\ldots,x^{(n)}_i) \) T, \( x^{(i)}_j \) represents the jth feature of the i-th sample, and the eigenvalue geometry of \( x^{(i)}_j \) is \( V=\{a_{j1},a_{j2},\ldots,a_{jw}\} \), where \( ALJ \) represents the n-th feature of the jth feature Eigenvalues, so the value ranges of \( j \) and \( l \) are \( j=1,2,\ldots,n, l=1,2,\ldots,w \), \( Y_i \in \{-1,1\} \), -1 describes alarm 1, 1 Describe alarm 2. Assuming that the training sample set contains t negative alarms one and m alarms 2, the empirical probability of alarm one and alarm two can be described by equations (4) and (5):

\[
p(Y = 1) = \frac{m}{N}
\]

\[
p(Y = -1) = \frac{t}{N}
\]
On the premise that alarm one and alarm two have been obtained, the conditional probability of \( x^{(j)} \) can be estimated using the maximum likelihood method, as shown in formulas (6) and (7).

\[
\begin{align*}
p(x^{(j)} = a_{jl} | Y = 1) &= \frac{\sum_{i=1}^{N} \text{count}(x^{(j)} = a_{jl}, y=1)}{m} \\
p(x^{(j)} = a_{jl} | Y = -1) &= \frac{\sum_{i=1}^{N} \text{count}(x^{(j)} = a_{jl}, y=-1)}{m}
\end{align*}
\]

(6) (7)

Among them, \( \text{count}(x^{(j)} = a_{jl}, y=1) \) represents the number of times the feature with the characteristic value of \( a_{jl} \) occurs in alarm 1, \( \text{count}(x^{(j)} = a_{jl}, y=-1) \) represents the characteristic value. It is the number of occurrences of a feature in alarm 2. Therefore, for an alarm information data set to be classified \( x = (x^{(1)}, x^{(2)}, ..., x^{(n)}) \), \( T \), it is divided into The probability of alarm one can be calculated by the formula (8):

\[
p(Y = 1 | x) = \frac{m}{N} \prod_{j=1}^{n} p(X^{(j)} = x^{(j)} | Y = 1)
\]

(8)

The probability of being classified as alarm two can be calculated by the formula (9):

\[
p(Y = 1 | x) = \frac{t}{N} \prod_{j=1}^{n} p(X^{(j)} = x^{(j)} | Y = 1)
\]

(9)

For a sample set of alarm information to be classified, the correct classification objective function can be calculated using formula (10):

\[
f(x) = \begin{cases} 
1, & \text{if } p(Y = 1 | x) > p(Y = -1 | x) \\
-1, & \text{if } p(Y = 1 | x) < p(Y = -1 | x) 
\end{cases}
\]

(10)

2.3. K nearest neighbour method

K nearest neighbour is a semi-supervised learning method. The K-nearest neighbour algorithm adopts the idea of statistical classification, which can be widely used in warning information classification. The formal description is as follows: Given a training set, for the alarm information sample to be classified, the K samples nearest to the relevant sample can be found in the training set, and the category of the sample to be classified can be judged as the category where the majority of the K samples are located [3]. The specific description of the K-nearest neighbour classifier is as follows: Suppose that the alarm information training set sample \( D = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\} \) contains \( N \) samples, where \( x_i \) represents the ith alarm The feature vector of the information, \( y_i \) represents the category of the alarm information, the classification process of the category of the feature vector is as follows:

- ① Calculate the distance \( d(x, x_i) \) between \( x \) and \( x_i \), \( i = \{1, 2, ..., n\} \);
- ② Sorting distance \( d(x, x_i) \) in ascending order, \( i = \{1, 2, ..., n\} \);
- ③ Select the first \( K \) samples with the smallest distance, and determine the category of each sample;
- ④ Count the number of occurrences of each category;
- ⑤ Divide the samples to be classified into the emotional category with the most frequent occurrences.

3. Mining and analysis of alarm information in the metering system
3.1. Design of alarm information mining algorithm for the metering system

Support Vector Machine (Support Vector Machine) is a supervised learning algorithm based on statistical learning theory. It was put forward by Vapnik et al. The subset and the discriminant function in the subset minimize the actual risk of the learning machine. Therefore, the slightest error classifier can be obtained by learning the training samples, and the error of the test case set is still minimized [4]. Specifically, the basic operating idea of the support vector machine is: in the case of linear separability, the optimal classification hyperplane of the two types of samples can be found in the original solution space; in the case of linear inseparability, the support vector machine can add slack variables to perform practical analysis and use non-linear mapping to map the samples in the low-dimensional input space to the high-dimensional attribute space and become linear, so that the support vector machine can learn and test in the high-dimensional classification algorithm, and can be obtained in the feature space The best classification hyperplane. Then, the classification hyperplane is optimized, and the structural risk minimization principle is adopted to construct the optimal classification hyperplane in the attribute space so that the classifier obtains the global optimum.

Figure 1. SVM Schematic diagram

Traditional support vector machines cannot make accurate judgments in the classification of alarm information in the measurement system, which reduces the recall rate of classification. In order to solve the above problems, this paper introduces the basic idea of regression prediction. The regression prediction idea can effectively use the algorithm to collect alarm information context data, fully consider the content of alarm information classification, and combine the current alarm information expression background, which can significantly improve the accuracy of alarm information analysis [5]. Therefore, this paper proposes an improved SVM alarm information classification algorithm, which is described in detail.

Assuming that the sample data \((x, y)\) collected by the alarm information classification follows the probability distribution \(P(x, y)\), then the expression of the regression prediction function can be set as shown in formula (11).

\[
F = \{f|f(x) = w^T\Phi(x) + b, w \in \mathbb{R}^n\}
\]
At the same time, the risk function is introduced in the execution of the SVM alarm information classification algorithm based on regression prediction, as shown in formula (12).

\[
R_{\text{reg}} = \frac{1}{2} \|w\|^2 + C \cdot R_{\text{emp}}[f] \tag{12}
\]

The SVM algorithm based on regression prediction proposed in this paper includes training and classification modules. The functions of these two modules can be described as follows:

1) SVM training module based on regression prediction.

In the training module of the SVM algorithm based on regression prediction, its key functional modules include four parts, namely the control module, the alarm information classification training module, the alarm information classification data reading module, and the alarm information classification evaluation module. In the algorithm of this paper, these four functional modules are effectively integrated to complete the evaluation of alarm information classification. The algorithm execution flow is shown in Figure 2.

**Figure 2. SVM training module algorithm execution process based on regression prediction**

1) First, determine the input alarm information feature serialization conditions according to the needs, determine the alarm information feature sequence, and count the value range of a context alarm information feature sequence.

2) Then call the alarm information vocabulary feature library, read the control function and classification algorithm evaluation function of this module, immediately perform the statistical work of the alarm information feature content sequence, and save the result of the alarm information feature evaluation value into the LIST structure data.

3) The algorithm takes out the data saved in the LIST, transmits it to the prediction model, trains the alarm information classification function module, and generates an alarm information classification model.

2) SVM prediction module based on regression prediction.

After training the alarm information classification model, use the following steps to classify the actual alarm information. The key steps are described as follows:
1) Obtain the actual data of the alarm information content, and set the value range of the alarm information feature context according to the requirements.
2) Statistical alarm information features appear in the context of the location.
3) Call the alarm information classification function in this module to evaluate the actual alarm information orientation. The specific process of the improved algorithm is shown in Figure 3.

![Diagram](image)

**Figure 3. SVM Prediction Module Algorithm Execution Process Based on Regression Prediction**

After the above process is executed, the training model of the improved SVM algorithm can be obtained and applied to the actual alarm information classification process. Detailed experiments and experimental results will be described in the following chapters.

### 3.2. Analysis of the experimental results of the alarm information mining algorithm of the metering system

This paper extracts relevant alarm data from the SCADA system, EMS system, OPEN-3000 management system, remote meter reading, and other metering systems. In order to verify the effectiveness of this algorithm, it is combined with linear regression algorithm, naive Bayes method and K nearest neighbours. The experimental results of the methods are compared in order to be able to highlight the effectiveness of the algorithm in this paper. First of all, this article identifies the types of warning information. The evaluation methods used are accuracy, recall rate, and F value. The operating results of the three identification methods are shown in Table 1.

| Algorithm                          | Precision | Recall |
|------------------------------------|-----------|--------|
| K nearest neighbour algorithm      | 0.61      | 0.72   |
| Naive Bayes Classification Algorithm | 0.64      | 0.71   |
| Linear regression algorithm        | 0.66      | 0.73   |
| SVM algorithm for regression prediction | 0.83      | 0.83   |

*Table 1. Comparison of the running results of the three algorithms*
The experimental results show that the SVM algorithm based on regression prediction proposed in this paper has higher accuracy, recall rate, and F value when identifying warning information. Good results.

4. Conclusion
With the application and management of computer technology in the operation of the power grid, the number and types of access devices in the power metering system are large, and the amount of alarm information data is significant. Manual analysis cannot locate information quickly and in real-time. This paper designs a support vector machine algorithm based on regression prediction, which can accurately classify alarm information, facilitate judgment of different types of alarm information, respond in time, and optimize grid operations.

References
[1] CHETLUR S, WOOLLEY C, VANDERMERSCH P, et al. CuDNN: ef-ficient primitives for deep learning [C]. Conference and Workshop on Neural Information Processing Systems NIPS, 2014: 11-19.
[2] ABADI, MARTÍN, BARHAM P, CHEN J, et al.TensorFlow: a systemfor large-scale machine learning [C]. OSDI'16: Proceedings of the12th USENIX conference on Operating Systems Design and Implement-tation November, 2016: 265-283.
[3] VASILACHE N, JOHNSON J, MATHIEU M, et al. Fast convolutionalcts with FBFFT: a GPU performance evaluation [C]. San Diego: International Conference on Learning Representations, 2015.
[4] Friedman N, Geiger D, Goldszmidt M. Bayesian network classifi- ers [J]. Machine Learning, 1997, 29( 2-3): 131-163.
[5] Chickering D M. Learning Bayesian networks is NP complete [C]. Fisher D, Lenz H, eds, Artificial Intelligence and Statistics V, Springer-Verlag, 1996: 121-130.
[6] LOW Y, GONZALEZ J E, KYROLA A, et al.GraphLab: a new framework for parallel machine learning [C]. The 26 Conference on Uncertainty in Artificial Intelligence (UAI2010), 2014: 1-11.
[7] Shuo Zhao. Research and application of cloud computing and machine learning algorithms in power load forecasting [D]. North China Electric Power University, 2014: 1-3.
[8] Shangzhi Yu. Application of machine learning algorithms in short-term power load forecasting [J]. Science Technology and Engineering, 2013, 13(08): 2231-2234.