Optimization research on short-term load forecasting method for electric vehicles based on SSA-SVM

Jiaqi Sun 1, a, Linlin Tan 2, b, Jinpeng Zhu 3, c, Xin Cheng 4, d, Xiaoqi Shen 2, e

1 Suzhou Joint Graduate School of Southeast University, Nanjing City, Jiangsu Province, China;
2 School of Electrical Engineering of Southeast University, Nanjing City, Jiangsu Province, China;
3 School of Civil Engineering and Architecture, Beijing Jiaotong University, Beijing, China;
4 School of Software, Southeast University, Nanjing, Jiangsu Province, China.

a jqjpwyykl@163.com, b tanlinlin@seu.edu.cn, c 1009326159@qq.com, d C18844209300@163.com, e 1015656463@qq.com

Abstract. In order to improve the accuracy of short-term electric vehicle load forecasting, a combined forecasting model based on Sparrow Search Algorithm (SSA) and Support Vector Machine (SVM) is constructed. Using SVM model to predict electric vehicle load, accelerating the convergence speed of the prediction model through sparrows search algorithm, conducting global optimization in the solution space, expressing the impact of calculated parameters on the model through fitness, searching for the optimal model parameter data, and improving the overall accuracy of the prediction model. Through simulation experiments, comparing the prediction results of SVM models, and SSA-SVM shows better performance in electric vehicle load prediction.

Keywords: Short term electric vehicle load forecasting; Sparrow Search Algorithm (SSA); Support Vector Machine (SVM); Backpropagation Neural Network (BPNN); Long Short Term Memory (LSTM); Intelligent optimization algorithms.

1. Introduction

In recent years, with the effective promotion of new energy electric vehicles, the random charging demand brought about by the large-scale growth of electric vehicles [1] has brought severe challenges to the stability and safety of the national power grid, so it is necessary to study the impact of large-scale electric vehicle charging on the power grid. At the same time, it is one of the current research hotspots and difficulties to effectively predict future load changes by analyzing historical charging load data [2], so as to cope with and formulate corresponding solutions.

At present, the methods for EV charging load prediction can be divided into two categories [3]: one is based on statistics [4], and the other is based on machine learning [5]. Compared to statistical methods, machine learning methods have better performance in data processing, data fitting, and model robustness. The support vector machine (SVM) algorithm mentioned in reference [6] is a binary classification algorithm in machine learning that processes classification, prediction, and other problems in a supervised learning manner. Its characteristic is that when the sample data is linearly indivisible, the non-linear mapping maps the samples from the low dimensional input space to the high-dimensional space by adding relaxation variables, making it linearly separable and searching for the optimal classification hyperplane. This model performs well in short-term electric vehicle load forecasting [7]. In practical load forecasting applications, load sample data has strong randomness, and there are cases of sample data loss or incorrect filling. This article proposes an improved SVM prediction model based on Sparrow Search Algorithm (SSA) [8] to address the issue of insufficient sample data and insufficient accuracy and accuracy caused by non-optimal model parameters in training models. Compared with optimization algorithms such as Genetic Algorithm [9], Cuckoo Search Algorithm [10], and Particle Swarm Optimization [11], The global optimization ability is strong, and the SVM model parameters can be calculated by using the fitness function to improve the prediction. According to the experimental results, the improved prediction model has higher prediction accuracy.
2. Introduction to Algorithm

2.1 SSA optimization algorithm

The principle of the sparrow algorithm refers to dividing the sparrow population into discoverers, joiners, and vigilantes. By comparing safety thresholds, warning values, and fitness, the positions of discoverers, joiners, and vigilantes are continuously updated to complete foraging behavior. Different types of sparrows correspond to a solution at an optimization problem, and through this foraging behavior, the global optimal solution is searched.

The location update description of the discoverer is as follows:

\[
X_{i,j}^{t+1} = \begin{cases} 
X_{i,j}^t \times \exp \left( \frac{-i}{a \times \text{iter}_{\max}} \right), & R_2 < ST \\
X_{i,j}^t + Q \times L, & R_2 \geq ST 
\end{cases}
\]

(2)

In the formula: \( t \) is the current number of iterations; The range of \( j \) is \( \{1,2,...,d\} \), where \( d \) is the dimension of the variable to be optimized; \( X_{i,j}^t \) Represents the value of the \( i \)-th sparrow at the \( j \)-th dimension position in the \( t \)-th iteration; \( \text{iter}_{\max} \) is the maximum number of iterations in the process; \( a \) It is a random number ranging from 1 to 0; \( R_2 \) represents the alarm value, \( ST \) represents the safety threshold; \( Q \) is a random number that satisfies a normal distribution; \( L \) is a 1xd matrix with all elements being 1. When \( R_2 < ST \), it indicates that the alarm threshold has not been reached and the environment is relatively safe; When \( R_2 > ST \), it indicates reaching the alarm threshold and being in a hazardous environment.

The updated description of the joining party's location is as follows:

\[
X_{i,j}^{t+1} = \begin{cases} 
Q \times \exp \left( \frac{x_{i,j}^t - x_{i,j}^w}{\beta} \right), & i \times \frac{n}{2} \\
X_{i,j}^t - X_{i,j}^{t+1} \times A \times L + X_{i,j}^{t+1}, & i \times \frac{n}{2} 
\end{cases}
\]

(3)

In the formula: \( X_{i,j}^{t+1} \) represents the \( t \) best position occupied by the discoverer in the \( (t+1) \)st iteration of the population; \( X_{i,j}^w \) The worst global position in the current situation; \( A \) represents a 1xd matrix where elements are randomly assigned as 1 or -1.

Guardians account for 10% -20% of the entire population, and in the sparrow population, \( f \) represents the sparrow’s fitness. The updated location description is as follows:

\[
X_{i,j}^{t+1} = \begin{cases} 
X_{i,j}^w + \beta \times \left| X_{i,j}^w - X_{i,j}^t \right|, & f_t \neq f_g \\
X_{i,j}^t + k \times \left( \frac{X_{i,j}^w - X_{i,j}^w}{f_t - f_g} + \epsilon \right), & f_t = f_g 
\end{cases}
\]

(4)

In the formula: \( X_{i,j}^w \) is the current global optimal position; \( \beta \) It is a step size control parameter; \( K \) is a random number that represents the direction of sparrow movement; \( f_t \) represents the fitness of the \( i \)-th sparrow; \( f_w \) represents the current global worst fitness value; \( f_g \) represents the current global best fitness value; \( \epsilon \) For a very small constant.

2.2 SVM algorithm

is a binary classification model with the largest spacing of linear classifiers in the feature space,. SVM method has a relatively fast convergence speed and better robustness in dealing with regression prediction problems.

Assuming the training dataset \( T \) for electric vehicle load is:

\[
T = \{(x_1,y_1), (x_2,y_2), ..., (x_m,y_m)\}
\]

(5)

In the formula: \( x_m \) the m-th input feature vector in the original sample, and \( y_m \) is the number of categories corresponding to \( x_m \). Due to the nonlinearity of input and output in most cases, it is necessary to establish a mapping function to map the sample space of the original input feature to a
higher-order feature space, and use the mapping function to correspond the original spatial feature to a high-dimensional space. The purpose of the SVM model is to find the optimal hyperplane by solving the following optimal programming equation:

$$\begin{align*}
\min P(\theta) &= \min \frac{1}{2} |\theta|^2 + c \sum_{i=1}^{n} e_i^2 \\
\text{s.t.} \quad \theta \phi(x_i) + b &\geq y_i - e_i, i = 1, 2, \ldots, m
\end{align*}$$  \(6\)

In the formula: \(f(w)\) is the predicted regression function; \(\theta\) is the normal vector of the hyperplane; \(b\) is the bias amount determined during the training sample process; \(c\) is the penalty coefficient for the error term, and \(c > 0\); \(e_i\) is the error value of the regression prediction model. Based on the above constraints, the equation can be solved by establishing a Lagrange equation, which is:

$$L(\theta, b, e, \mu) = \frac{1}{2} |\theta|^2 + c \sum_{i=1}^{n} e_i^2 - \sum_{i=1}^{n} \xi_i (\theta \phi(x_i) + e_i + b - y_i)$$  \(7\)

In the formula: \(\phi(x)\) is the mapping relationship function; \(\xi_i\) represents the Lagrangian of the sample, based on the first-order optimal condition, the variables in the equation are evaluated as first-order partial derivatives one by one, and the equation is solved to obtain the hyperplane optimal solution. The RBF kernel function used in this article is expressed as:

$$K(x_i, x_j) = \exp(-\eta |x_i - x_j|^2)$$  \(8\)

In the formula: \(\eta > 0\) is an adjustable parameter, and \(x_j\) is the specified center point.

Based on the above results, the SVM predictive regression model constructed in this article is:

$$f(x) = \sum_{i=1}^{n} (\xi_i K(x_i, x)) + b$$  \(9\)

### 3. Design of Short term Electric Vehicle Load Forecasting Algorithm Based on SSA-SVM

In SVM prediction models, there are currently problems such as long training time and poor data coupling, partly due to the significant impact of regularization penalty coefficient \(c\) and kernel parameters on prediction results. The larger the penalty coefficient \(c\) value, the lower the model's acceptance of prediction errors, which may lead to overfitting; A smaller value indicates a higher acceptance of the prediction error, which may lead to underfitting. A penalty coefficient \(c\) that is too large or too small can have a certain impact on the fitting effect.

By adding an RBF kernel function, the \(\eta\) parameters in the kernel function can map nonlinear data to a higher dimensional feature space. The \(\eta\) larger the value, the fewer support vectors there are; conversely, the more support vectors there are. Through the analysis and use of SSA algorithm optimization, it can be concluded that SVM models equipped with SSA optimization algorithm can achieve faster convergence speed during model training and accelerate the calculation speed of SVM. After establishing the SVM model, analyze its model requirements, determine the fitness function that SSA needs to calculate, and quickly find the optimal parameters for the penalty coefficient \(c\) and kernel parameters \(\eta\) of the SVM model to improve the accuracy of model prediction.

According to the demand for prediction results, the RMS (root mean square) of the predicted and true values of electric vehicle loads is used in this model to express the individual fitness function. The reason is that the lower the RMS of the predicted and true values \(\eta\), the more accurate the model is. The SSA algorithm searches for the optimal solution in the global fitness function and optimizes the penalty coefficient \(c\) and kernel function parameters of SVM through the SSA algorithm. The process is to initialize the SSA-SVM model parameters; By continuously updating the position information of the joiners, discoverers, and vigilantes through individual fitness
functions, the position of sparrows is a function value; Determine the optimal function value after searching for the optimal sparrow position; The prediction process is shown in Figure 1, and the specific steps are as follows:

![SSA-SVM Model Prediction Flow Chart](image)

Fig. 1 SSA-SVM Model Prediction Flow Chart

1. To build an SSA optimization algorithm, first set the initialization parameters and set the warning value, ranging from 0.5 to 1; Set the dim parameter, dim is the number of decision variables, set to 2 in this model; In this model, the number of sparrows is set to 100, indicating a random distribution of 100 solutions on the fitness function. The number of discoverers and joiners is 70% and 30% of the number of sparrows, respectively. The alerter is 20% of the number of sparrows, and the alerter is to ensure that the sparrow population is within a safe threshold range. It can be understood that the searched function value is within a limited range.

2. By describing the SSA algorithm, determine the fitness function of the sparrow search algorithm, which is the RMS of the predicted value and the true value. In each iteration process, by setting the number of iterations to find the global or local optimal function value, generally speaking, after the iteration, obtain the optimal sparrow individual position based on the fitness function; Obtain the \( \eta \) optimal value of model parameters \( c \) and \( d \).

3. Determine whether the termination condition. If it is met, the algorithm ends. Otherwise, update the location information of the discoverer and joiner again. Assign the optimal parameter \( c \) and \( \eta \) to the support vector machine. The predicted value of the SSA-SVM model is output, and the prediction result of the support vector machine before optimization is compared.

4. Analysis of 3 examples

Taking the historical load data of electric vehicles in a certain area of Suzhou City, Jiangsu Province in the first quarter of 2020 as an example data sample, the sampling interval is two hours. After data processing, the original samples were used as experimental samples for a total of 1080 experimental data, with 70% being the training set and 30% being the testing set. By using the SSA-SVM model and SVM, BP [12], LSTM [13] models to predict the load data of electric vehicles, the accuracy and error of data prediction between each model are compared. There are missing values and outliers in the original sample data.

After model training, the partial prediction results of its SSA-SVM, LSTM, SVM and BP models are compared with the sample data in Figures 2, 3, 4 and 5 respectively. It can be concluded that compared with other models, the SSA-SVM model proposed in this paper has a higher degree of data coupling and more accurate prediction results.
To make the experimental results more intuitive, three performance indicators, RMSE, MAE, and MSE, are used in this article to demonstrate the accuracy of different prediction models in predicting data. Table 1 shows the comparison results of the four models under the three performance indicators. Compared with the SVM model, the SSA-SVM model has reduced RMSE, MAE and MSE to a certain extent, indicating that the SSA-SVM model has improved the accuracy of electric vehicle load forecasting to a certain extent.

Table 1. Comparison of prediction performance indicators of four models

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>RMSE</th>
<th>MAE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP</td>
<td>0.0203</td>
<td>3.047</td>
<td>13.612</td>
</tr>
<tr>
<td>LSTM</td>
<td>0.0132</td>
<td>0.037</td>
<td>1.473</td>
</tr>
<tr>
<td>SVM</td>
<td>0.0149</td>
<td>0.031</td>
<td>0.967</td>
</tr>
<tr>
<td>SSA-SVM</td>
<td>0.0103</td>
<td>0.028</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Compared to the three prediction models of BP, LSTM, and SVM, the RMSE of SVM has decreased by a maximum of 49.2%; The MAE of the SVM model reaches a minimum of 0.028; For the MSE indicator, the SVM model is only 0.023, which indicates a high degree of coupling between predicted and true values, and the accuracy of the model's prediction results is high.

5. Conclusion

The paper proposes an SSA-SVM electric vehicle load prediction algorithm, which optimizes SVM using intelligent optimization algorithm SSA. A SSA-SVM electric vehicle load prediction simulation model is established, and compared with LSTM, BP, and SVM prediction models. The results show that using SSA-SVM optimization algorithm significantly improves the parameter optimization of SVM model, accelerates model convergence and calculation speed, and significantly improves prediction accuracy. The model can have a higher degree of coupling data and more accurate prediction.

This work was supported in part by the Qing Lan Project of Jiangsu and Project of Jiangsu Provincial Key Laboratory of Smart Grid Technology and Equipment.
References


