International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 4 ǁ 2022 ǁ PP. 61-70

Prophet-LSTM Combination Model Based on PSO

Bowen Hu[1], Jun Li[1*]
College of Compuer Science & Technology, Qingdao University, Qingdao 266071, China

Abstract
With the development of the Internet and e-commerce, China has become the world's largest online sales market, online shopping this consumption pattern has gradually become an indispensable way of life in our national daily life. As an important part of enterprise management, sales forecasting can make merchants more accurate in sales, avoid product overstocking, reasonably manage product inventory, arrange production schedules, timely supplement products, and avoid the phenomenon of short supply. Compared with traditional offline store sales, e-commerce has advantages such as easy data collection, rapid data processing, and huge data volume, which can be used to forecast commodity sales more conveniently and accurately. In this paper, the Prophet model is combined with the multi-layer LSTM model. Finally, the PSO algorithm is used to combine the weights of the prediction results of the two models, so as to obtain a sales prediction network model that is more suitable for the sales fluctuation characteristics of e-commerce sales before and after the activity day. In addition, the sales data set of clothing e-commerce on Taobao was used to conduct a comparative test. The single Prophet model, multi-layer LSTM model, and traditional ARIMA time series regression model were compared. The experimental results show that in the case of e-commerce with large sales fluctuations, the performance of the Prophet-LSTM-PSO combined model is better, and the prediction accuracy is further improved.

Key word*：*Sales forecast; LSTM; PSO; Prophet;

Date of Submission: 05-04-2022 Date of acceptance: 20-04-2022

In recent years, the development of e-commerce in my country has shown a surge, and the competition among e-commerce companies has become increasingly fierce. At the same time, online sales are more complex and multilateral than traditional offline retail. , In today's era of booming e-commerce companies, is also full of new challenges. At this point, sales forecasting has become an important indicator for e-commerce companies to control inventory management and change their marketing strategies.

At present, commodity sales forecasting is mainly divided into methods based on statistics, methods based on machine learning, methods based on data mining, and methods based on deep learning [1]. Sales forecasting has been studied by scholars at home and abroad for many years, and many models have been proposed to improve the accuracy of forecasting [2]. In 2017, Facebook developed a time series forecasting algorithm Prophet [3], which can fit the curve of time series data, and can also analyze the impact of holidays and trend change points, so if the experimental conditions are the same, this method can The predictive power shown is better than the general model. Na Ge [4] combined Prophet with long short-term memory neural network, and the proposed product sales prediction model outperformed the single model. However, the optimal weight ratio is not obtained. Then, according to the characteristics of the two components, the Prophet model and Gaussian process regression are respectively applied to predict the two components. According to the experimental data, it can be seen that the model he constructed is superior to the traditional time series forecasting method [5]. K. Krishna Rani Samal et al used SARIMA (Seasonal Autoregressive Integrated Moving Average) and Prophet model to predict air pollution levels in Bhubaneswar. The urban PM2.5 concentration forecast is roughly predicted by the combination of the SARIMA seasonal difference autoregressive moving average model and the Prophet model. The model mainly uses seasonality as an influencing factor, so the accuracy of the forecast results is general[6]. Wu [7] combined particle swarm optimization and a support vector machine to predict car sales. Du Nannan et al. [8] used the LAE algorithm to combine the Prophet model with the LSTM model to predict the water level change of Honghu Lake. The proposed Prophet-LSTM model could not only use the LEA weight allocation method, and there was no corresponding weight comparison experiment. Many combined models use neural network models, and at the same time cite the advantages of other models, which can improve the robustness and accuracy of predictions. The data obtained by multiple models can be integrated into a whole, which can also effectively reduce the estimation error.

Based on PSO (particle swarm optimization algorithm), combined with the combined model of Prophet (prophet model) and LSTM neural network model, this paper proposes an e-commerce sales forecast model.

Because Prophet can adapt to the trend points of data changes in time series forecasting, and LSTM neural network can mine nonlinear and aperiodic features hidden in long periods [9]. The combined model can take advantage of the advantages of multiple models, make up for the shortcomings of a single model, and improve the prediction accuracy of the model. Finally, the PSO algorithm is used to solve the combined coefficients of the prediction results of the two models. Get the final sales forecast data. And compared with the single model Prophet model, LSTM model, and traditional ARIMA statistical model. The results show that the proposed PSO-Prophet-LSTM combined model outperforms the independent forecasting model in sales forecasting accuracy, and its forecasting performance is greatly improved.

I. Model Algorithm Design

1.1 Prophet prediction model

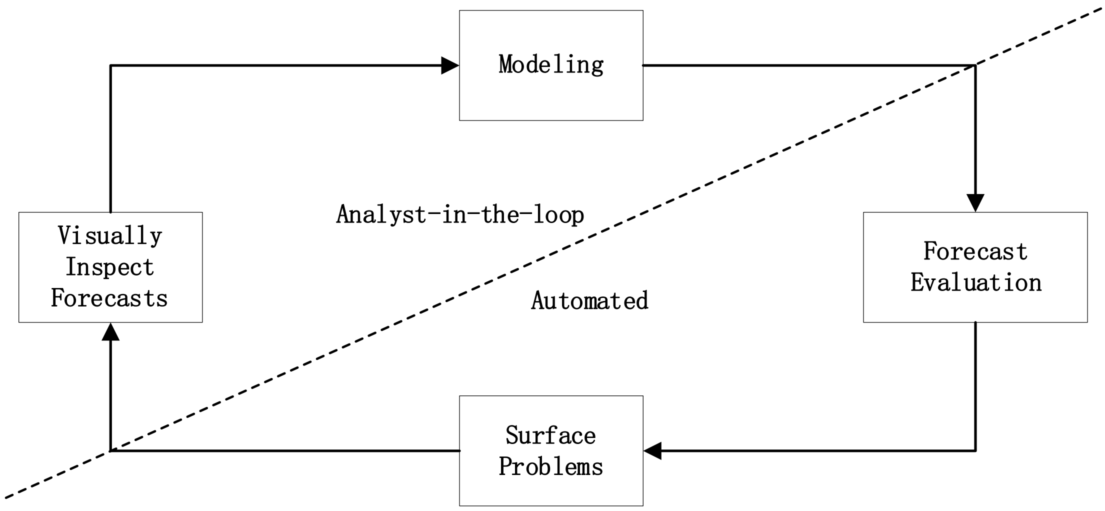


Figure 1. Prophet model program.

Prophet [10] is an open source software developed by Facebook developers. It is a predictive time series model based on an additive model that fits non-linear trends as well as annual, weekly, and daily seasonality and holiday effects. It performs well in time series data with strong seasonal effects and historical data with multiple seasons, and generally handles outliers well [11]. The basic flowchart of the Prophet model is shown in Figure 1 below.

As can be seen from the above figure, the Prophet model consists of several modules, namely model construction, problem presentation, prediction evaluation, and visualization of prediction results and analysis. According to the above 4 modules, the prediction steps of the Prophet model are as follows:

1. Modeling: Modeling of time series. This process is mainly based on the background conditions of the business to select an appropriate prediction model.
2. Forecast Evaluation: Forecast evaluation stage. Perform small-scale data verification operations, use business analysis and actual data to continuously optimize model parameters, and then use scientific standards to judge the quality of the model.
3. Surface Problems: Presenting the problem stage. Present all the influencing factors of the model, and then carry out the cleaning of these factors.
4. Visually Inspect Forecasts: Visually present forecast results, and then perform legend research. Use the prediction results to analyze the subsequent improvement direction of the model.

The above four parts are connected in an end-to-end fashion and divided into two parts, modeling and evaluation, through the diagonal line in the figure. Through such a closed loop, the model combines modeling and evaluation to further complete the rapid iterative optimization of time series models. The prophet is a time series model with decomposable characteristics. The whole model consists of three parts: growth (growth trend), seasonality (seasonal trend), and holidays (the impact of holidays on the predicted value). The combined formula is:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \#(1)$$

(1) $g(t)$ is a trend function that simulates aperiodic changes in a time series. It consists of two optional functions: piecewise linear function and logistic regression function. Among them, the logistic regression function is generally used to calculate the trend item. The basic form is as follows:

$$\mathrm{g}(t) = \frac{C(t)}{1 + e^{\left(-\left(k+\alpha(t)^T\delta\right)\left(t-\left(m+\alpha(t)^T\gamma\right)\right)\right)}} \#(2)$$

$$\alpha(t) = \alpha_j(t) = \begin{cases} 1, t \geq S_j \\ 0, \text{otherwise} \end{cases} \#(3)$$

$$\gamma_{\mathrm{j}} = \left(s_j - m - \sum_{l<j} \gamma l\right) \cdot \left(1 - \frac{k + \sum_{kj} \delta_l}{k + \sum_{l \leq j} \delta_l}\right) \#(4)$$

In the above formula, $C(t)$ is the expected capacity system, k is the growth rate, m is the deviation parameter, s_j indicates the time stamp of the change point with the chronological change point j, $1 \leq j \leq s$, δ_j, δ_j is the speed change occurs At time δ_j, γ represents the smooth offset time δ_j, and $\alpha_j(t)^T$ represents the transposed vector $\alpha_j(t)$.

(2) $S(t)$ is the period term, and the period model is constructed through the Fourier series and P is the regular period of the time series ($P = 365.25$ when the day is the scale, and $P = 7$ when the week is the period). The specific form is as follows:

$$s(t) = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right)\right) \#(5)$$

(3) $h(t)$ is a non-periodic holiday item, but its impact is predictable to a certain extent. The impact of a particular holiday on the time series is about the same every year, so it is very necessary to include it in the forecast. Holidays are considered independent of each other, and $h(t)$ is of the form:

$$h(t) = Z(t)k\#(6)$$

$$Z(t) = [1(t \in D_1), \ldots, 1(t \in D_L)]\#(7)$$

$$k \in \mathrm{Normal}(0, v^2) \#(8)$$

Suppose i represents the ith holiday, then D_i represents the duration of the ith holiday, k represents the range of influence of the holiday, and k is affected by v. The larger the v, the greater the influence of the holiday trend on the model predicted trend.

(4) ε_t represents the interference term, which represents the special changes that the model cannot adapt to, and is generally considered to obey the Gaussian distribution.

1.2 Multilayer LSTM Prediction Model

The multi-layer LSTM model is used to build the LSTM model part, the three-layer LSTM neural network model is interspersed with the Dropout pooling layer, and finally, the Desen layer is used to output the prediction results.

1.3 Calculation of combined model weights

Combination prediction models generally have the advantages of a single model, and the system is affected by each internal model, according to different weight values, to increase the accuracy of the prediction results [12]. The scope of use and advantages of each model are different, and the focus of its information is also different. Therefore, it is difficult to achieve satisfactory results by using only one model to predict the data. However, after combining the models, it is necessary to consider the weight of each model in the prediction result. For multiple regression problems, the regression parameters can generally be calculated by gradient descent method, least squares method, and Newton method. However, the above calculation methods all have complex matrix operations and the solution process is long. The PSO algorithm is an intelligent simulation algorithm based on the rule of law for bird predation activities. Compared with the ant colony algorithm (ACO), the genetic algorithm (GA), etc., its advantages are that there are not many parameters, high efficiency and simple and easy to operate, fast convergence speed, etc. Therefore, the author uses the particle swarm algorithm to calculate the proportion of each model.

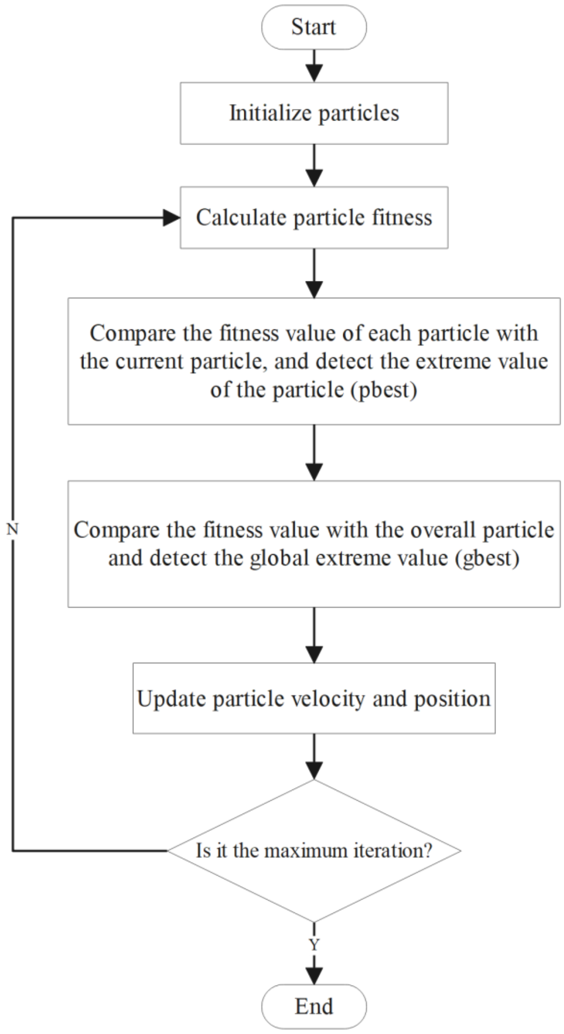


Figure 2. PSO algorithm flow chart.

The particle swarm optimization (PSO) algorithm was proposed by Kennedy and Eberhart (Kennedy and Eberhart, 1995). The basic idea of the PSO algorithm is to find the optimal solution through cooperation and information sharing among individuals in the group. In the PSO algorithm, a particle has its velocity and position. Velocity represents how fast it moves; position represents the direction of movement. Each particle in the D-dimensional space searches for the optimal solution and records it as the current individual extreme value (Particle best, pbest). Next, the algorithm compares all individual extrema in a particle swarm, and the best one is denoted as the global extremum (global best, gbest) of the entire particle swarm. All particles in the particle swarm adjust their velocities and positions via pbest and gbest. The particle swarm iterates the above process continues until it finds the optimal solution. The flowchart of the PSO algorithm is shown in Figure 2.

Suppose there are N particle swarms in the D-dimensional search space. The spatial position of the ith particle is expressed as $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$. The velocity of the ith particle is expressed as $v_i = (v_{i1}, v_{i2}, \dots, v_{id})$. The expression corresponding to the optimal solution of the ith particle in the whole search process is pbest$_i = (p_{i1}, p_{i2}, \dots, p_{iD})$. The expression of the global extremum optimal solution is: gbest $= (p_{g1}, p_{g2}, \dots, p_{gD})$. Particles update their velocity and position according to the following formula:

$$v_{id}^k = wv_{id}^{k-1} + c_1 r_1 (pbest_{id} - x_{id}^{k-1}) + c_2 r_2 (gbest_d - x_{id}^{k-1}) \#(9)$$

$$x_{id}^k = x_{id}^{k-1} + v_{id}^{k-1} \#(10)$$

where k is the number of iterations. c_1 and c_2 are acceleration constants used to adjust the maximum learning step size. W is the inertia factor, which is used to adjust the search range of the solution space, and r_1 and r_2 are uniform random numbers in the range of [0,1] to increase the randomness of the search.

Equation (9) consists of three parts: the first part is inertia or momentum, which reflects the movement habit of particles and represents the tendency of particles to maintain their historical speed. The second part is cognition, which reflects the particle's memory or memory of its previous experience and represents the particle's tendency to approach its historical best position. The third part is the social part, which shows the group's historical experience of particle information sharing and synergy, that is, how close the particle is to the historical optimal position of the neighborhood or population.

The calculation process of the PSO algorithm is as follows:

Step 1. Initialize the particle swarm (swarm size $\tilde{n}$), including random positions and velocities.

Step 2. Evaluate the fitness of each particle according to the fitness function.

Step 3. First, calculate the values of pbest and gbest. The values of pbest and gbest for the entire population are determined by comparing the corresponding fitness values. A single value pbest represents the best position of the particle in one iteration. Next, the algorithm compares the previous pbest with the current position and uses the better fitness function of the two as the new pbest. Then, the algorithm compares the fitness values of all particle positions with the global best position. Once the current fitness value is better than the global optimum, gbest will update the particle's pbest, and the calculation equation is:

$$\text{pbes } t_i(k) = \begin{cases} \text{pbes } t_i(k-1) & \text{when } f\big(x_i(k)\big) \geqslant \text{ pbest }_i(k-1) \\ & x_i(k) \text{ other} \end{cases} \#(11)$$

$$\text{gbes } t_i(k) = min\{f(p \text{ best }_i(k))\}\#(12)$$

Step 4. Update velocity and position according to equations (9) and (10).

Step 5. If the maximum number of iterations is satisfied, take the extreme value of the particle swarm as the optimal solution. Otherwise, go back to step 2 and start the next iteration.

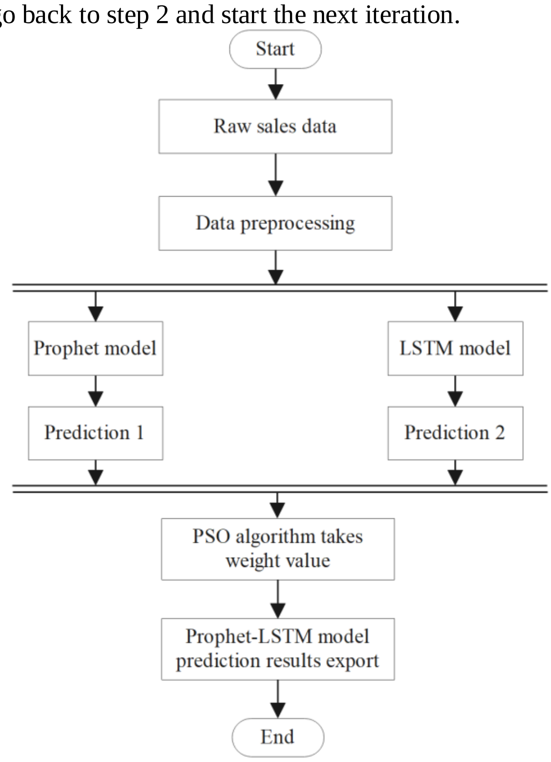


Figure 3. Basic flow chart of the combined model.

1.4 LSTM-Prophet Combination Model

In this paper, the advantages of the LSTM model and the Prophet model are fully considered, and the advantages of the Prophet model for predicting periodic time series and the advantages of the LSTM model for mining long-term historical data are combined to construct the LSTM-Prophet model. First, the LSTM model and Prophet model with higher prediction accuracy are constructed using historical sales data. The predicted values of Prophet and LSTM at time t are $\bar{P}_t$ and $\bar{L}_t$, respectively. Let the weight of Prophet be w_1 and the weight of LSTM be w_2. Finally, the best weight is obtained by the POS algorithm. The w_1 and w_2 obtained from this solution are the values that make the combined prediction model the closest to the true value. The combined weights of the combined model are calculated as follows:

$$Y(t) = w_1\bar{P}_t + w_2\bar{L}_t\#(13)$$

The basic flow chart of the combined model is shown in Figure 3:

II. Simulation experiment analysis

2.1 Dataset introduction

The data set used in this chapter is the Taobao real sales data set of a Tianjin clothing e-commerce company from January 1, 2017, to August 1, 2019. The clothing industry is greatly affected by seasonality. The sales data format of each clothing is as follows Table 1 shows:

Table 1. shows some datasets.

SKU	Category	Season	Tag_price	Date	Act_price	Num	…
S21380501	Sweater	Spring	399	20170307	279	1	…
S18462409	Shirt	Spring	269	20170307	188	1	…
S22123515	Jacket	Autumn	499	20170926	399	1	…
S18481888	Coat	Winter	1199	20180208	839	1	…
S19181824	T-Shirt	Summer	199	20180723	179	2	…
S18361863	Shirt	Spring	199	20190307	199	1	…
S19480878	Jacket	Spring	699	20190416	559	1	…

Through observation, we can see that each piece of data mainly includes attributes such as the SKU code to identify the style name, the category of the clothing, the season of the clothing listing, the price of the product, the date of sale of the product, and the price after the event discount.

2.2 Experimental data analysis and preprocessing

The original data will generate dirty data that is not conducive to model processing during the record storage process. Therefore, the original data needs to be further cleaned. For the data missing in the original data set, the missing values are necessary fields such as SKU code, sales data, etc. are deleted. For this piece of data, when the missing field is the listing season, the sales price, product type, etc. can be determined according to the SKU code, and the field can be completed according to the unique SKU code of the product.

Preliminarily clean the data and then encode some abnormal data with One-Hot keywords according to the scenarios in the data mining process, such as 618, the e-commerce event day, and the dates when the sales of holiday commodities such as Double Eleven surge. E-commerce merchants will have consumption activities organized by the platform to stimulate consumption at intervals, resume holiday activity dates according to the annual activity date, and map them one by one with the sales data. And in observing the e-commerce sales data, it is found that there will be a period before and after the event day that the sales of goods will be greatly reduced, and there will be two troughs. Therefore, a window period of 3 days before and after the event day is set. And set the Activity day parameter during the activity day to adjust the impact of fluctuations in product sales before and after the activity day. Some dates of e-commerce activity days and holidays are shown in Table 2 below:

Table 2. E-commerce activity dates.

Holiday/Event Day	Date	Duration
Spring Festival	2017-01-27,2018-02-14,2019-02-04	7
Goddess Festival	2017-03-07,2018-03-07,2019-03-07	3
Dragon Boat Festival	2017-05-30,2018-06-18,2019-06-07	3
618 Mid-Year Shopping Festival	2017-06-18,2018-06-18,2019-06-18	3
Mid-Autumn Festival	2017-10-24,2018-09-24,2019-09-13	3
National Day	2017-10-01,2018-10-01,2019-10-01	7
Double Eleven	2017-11-11,2018-11-11,2019-11-11	7

There are obvious periodic fluctuations in the processed data. To ensure that the time series can remain stable and the data can be approximately normally distributed, various methods are used to preprocess the time series data to make the data standardized and smoother [13]. Moreover, the better the stability of the time series, the more it satisfies the normal distribution law, its applicability is better, and the modeling accuracy is higher [14]. Therefore, a variety of transformation methods can be used to expand data processing. In this chapter, the normalization method is applied to normalize the data to facilitate the parameter learning of the subsequent model.

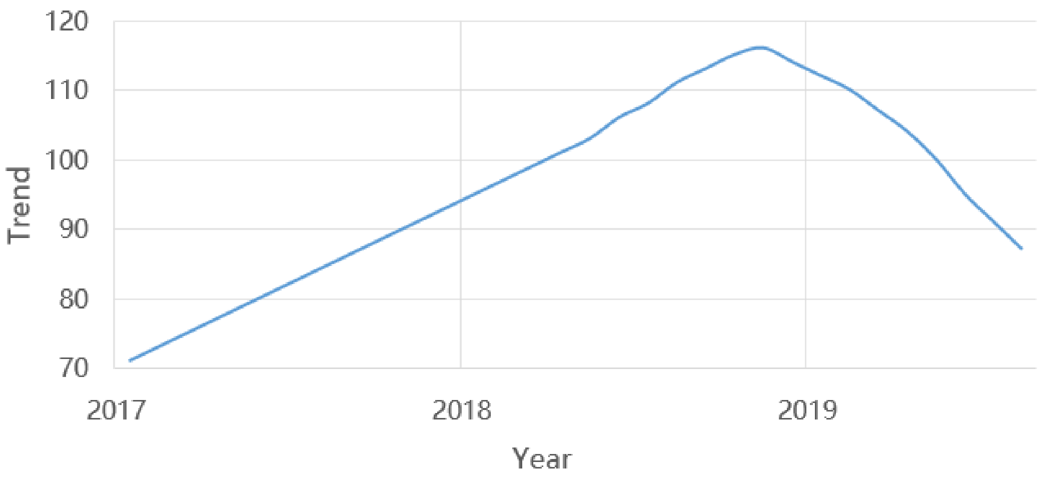
III. Model training results

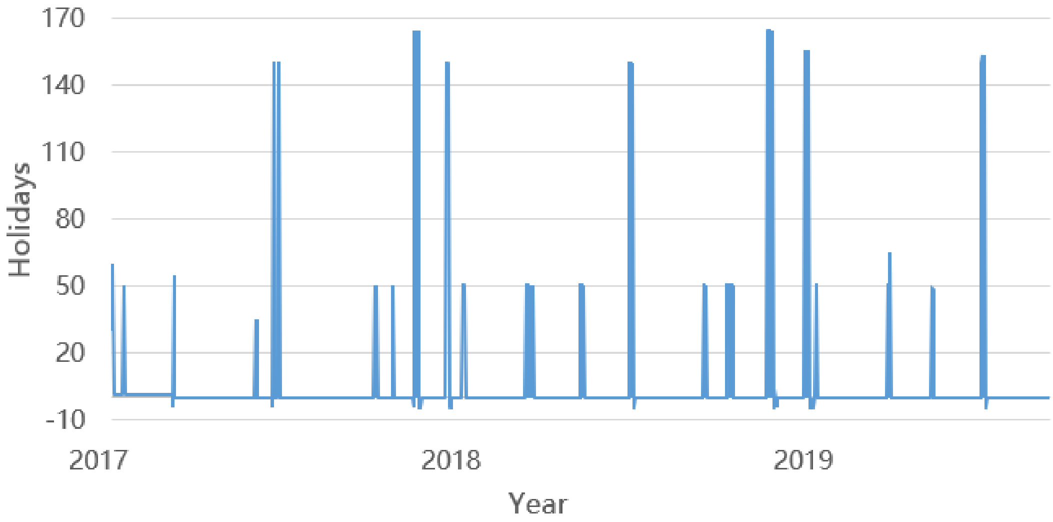
According to the 8:2 standard, all data is divided into a test set and a training set, and then the test set is processed by the Prophet model to complete the training operation. The model parameter settings are shown in Table 3 below:

Table 3. Prophet model parameter settings.

Parameter	Illustrate	Numerical Value
Changepoint_prior_scale	Control knee data	0.05
Interval_width	Confidence	0.8
Growth	Specify forecast type	Linear
MCMC_samples	Markov Monte Carlo Sampling	1000
Changepoint_range	Look for unusual data	0.9

The result data column of the model training contains the predicted values of individual components such as trend items, seasonal items, and event day items, as well as its prediction range. Through the prediction results of these individual data items, the composition of the prediction results can be observed. Plot the data into a visual analysis diagram as shown in Figure 4 below:





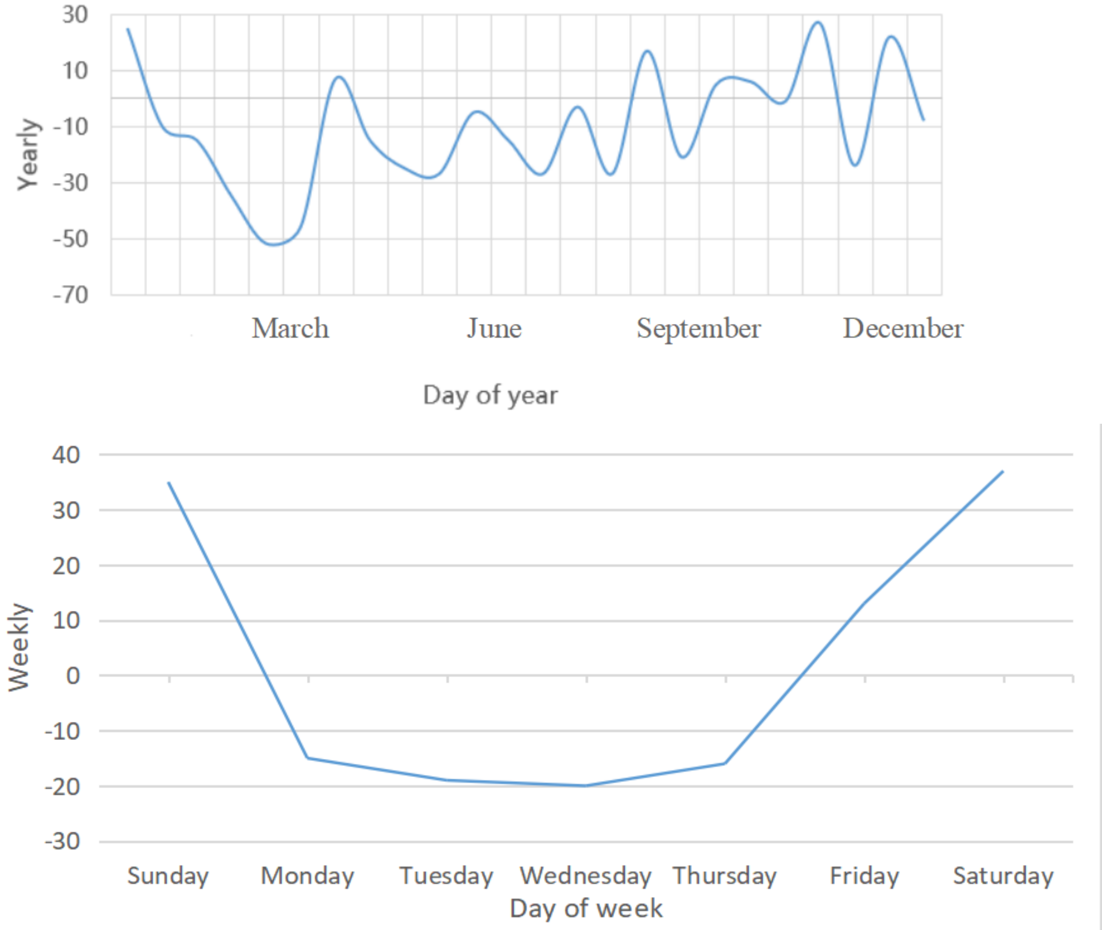


Figure 4. Composition analysis.

From top to bottom, the visual analysis graph shows the fitting situation of the trend item (trend), the holiday item (holidays), and the seasonal item (yearly, weekly). According to the above image, there is a trend of accelerated growth from the beginning of 2017 to the end of 2018, and the growth trend begins to decline in early 2019. The holiday chart shows that there will be a significant increase in sales during the e-commerce event day, which is in line with the general logic of e-commerce sales. Finally, there is the seasonal item. Through the observation of the experimental result graph, we can know that the sales volume of stores will decrease at the beginning of the year generally after the Spring Festival. In June and November, when 618 and Double Eleven are located There will be a peak, and during the Goddess Festival on March 7th and the start of school on September 9th, when Taobao has a large activity, its sales will increase significantly. From the weekly season chart, the sales on weekdays are relatively poor, but on Friday the sales increase and the increase is more significant on weekends, which is very similar to the actual situation. Finally, the predicted value through the Prophet model is $\bar{P}_t$.

Then, the processed data set is trained by the above-mentioned multi-layer LSTM model, and the final predicted value $\bar{L}_t$ is obtained. The optimization method adopted by the model is Adam, and the learning rate is set to 0.01.

Finally, the weight value is obtained by the PSO particle algorithm. If at time t, $P(t)$ is used to represent the predicted sequence value of the Prophet model, and $L(t)$ is used to represent the predicted sequence value of the LSTM model, and their corresponding weights are in order. After the combination of the Prophet-LSTM prediction model is completed, its calculation The equation is:

Finally, the weight value is obtained by the PSO particle algorithm. If it is at time t, $P(t)$ is used to represent the predicted sequence value of the Prophet model, and $L(t)$ is used to represent the predicted sequence value of the LSTM model. Their corresponding weights are W_1, W_2, and after the combination of the Prophet-LSTM prediction model is completed, its calculation equation is:

$$Y(t) = w_1 P(t) + w_2 L(t), w_1 + w_2 = 1, t = 1,2, \dots, N\#(14)$$

According to the above equation, the predicted sequence value $P(t)$ of Prophet and the predicted sequence value $L(t)$ of the LSTM model jointly determine the prediction result $Y(t)$ at time t according to the relevant weights. Use the PSO algorithm to calculate the weights of the combined model:

$$Y(t) = 0.65745P(t) + 0.34255L(t)\#(15)$$

To verify the superiority of the PSO algorithm in calculating the weight value, three weight-solving methods were selected for comparison. The detailed calculation content is as follows:

(1) According to the arithmetic average method, the weights of the two models of Prophet and LSTM are the same. In this way, if m single models constitute this combined model, and the weights of each model are equal, there is $\frac{1}{m}$. Then the calculation equation of the model is:

$$f_i = \frac{1}{m}\sum_{i=1}^{m} f_{it}, t = 1,2, \dots, n\#(16)$$

The weights calculated by the arithmetic mean method are as follows:

(2) The reciprocal difference rule pays more attention to the sum of squares of errors in the prediction results of the model. If the sum of squares of errors of the i-th model is very low, it means that the model has better prediction accuracy, and its weight is larger. If the sum of squared errors of the ith model is $S_i = \sum_{i=1}^{n}(y_i - f_i)^2$, f_i represents the predicted sequence value at time t, and y_i represents the true value at time t, then the weight calculation equation for this model is:

$$w_i = \frac{s_i^{-1}}{\sum_{i=1}^{m} s_i^{-1}}\#(18)$$

The weights calculated by the variance method are as follows:

$$Y(t) = 0.63P(t) + 0.37L(t)\#(19)$$

(3) The entropy weight method is more objective than the above two methods. In the entropy weight method, the prediction error value of each model is processed by normalization, and there are:

$$x_{it} = \frac{\varepsilon_{it} - \varepsilon_{imin}}{\varepsilon_{imax} - \varepsilon_{imin}}\#(20)$$

In the above equation, ε_{imin} and ε_{imax} are the maximum and minimum values of the ith model error sequence, and ε_{it} is the prediction error. Then at time t, find the model error value, the ratio p_{it} relative to the sum of all model error values, see the following calculation equation for details:

$$p_{it} = \frac{\varepsilon_{it}}{\sum_{i=1}^{n} \varepsilon_{it}}\#(21)$$

Then solve the entropy value S_i of the prediction error of the i-th model:

$$S_i = -K\sum_{j=1}^{n} p_{ij}\ln\, p_{ij}, K = (\ln\, n)^{-1}, 0 \le e_j \le 1\#(22)$$

Calculate the variance coefficient δ_i of the prediction error of the i-th model:

$$\delta_i = 1 - S_i\#(23)$$

Thus, the weight w_i of the i-th single-item model is obtained:

$$w_i = \frac{\delta_i}{\sum_{i=1}^{m} \delta_i}\#(24)$$

The entropy weight method can finally calculate the weights of the two models to be 0.42 and 0.58, respectively. At this time, the combined results are as follows:

$$Y(t) = 0.42P(t) + 0.58L(t)\#(25)$$

Table 4. Evaluation of combination algorithms.

Method	MAE	MAPE	RMSE
Arithmetic Mean	0.04237	0.03683	0.06752
Inverse Variance Method	0.03711	0.04604	0.04673
Entropy Weight Method	0.05561	0.04295	0.05276
PSO Algorithm	0.02549	0.03435	0.04016

The model evaluation indicators of the weighted combination model using the above four methods are calculated respectively, and the results are shown in Table 4:

In this chapter, MAE, MAPE, and RMSE are still used as the performance indicators of the model. The Prophet-LSTM model weighted by the PSO algorithm is compared with the ARIMA statistical model, the single-term Prophet model, and the LSTM model. The comparative analysis of the prediction results is shown in Table 5:

Table 5. Evaluation of prediction results.

Model	MAE	MAPE	RMSE
ARIMA	0.06189	0.08167	0.10238
Prophet	0.05710	0.04121	0.05703
LSTM	0.03284	0.03194	0.05186
PSO-Prophet-LSTM	0.02549	0.03435	0.04016

According to the above table, compared with the original time series ARIMA model, the advantages of the Prophet and LSTM models are more significant. However, compared with the two independent models, the combined Prophet-LSTM model has greatly enhanced prediction effect, and reduced the root mean square error and square absolute error of the model, compared with the independent model, the prediction accuracy of the combined model is better, and it is revealed that the combined model has a better prediction effect. It has the advantages of two independent models. Therefore, in this paper, the author chooses this combined model for research.

IV. Conclusion

In this paper, the advantages of the PSO-Prophet-LSTM combined model are expounded, the Prophet model and the calculation process of the PSO algorithm are introduced in detail, and then the combined model Prophet-LSTM model is constructed. Based on the multi-layer LSTM model, it is combined with the time series model Prophet, and the weight is calculated by the arithmetic average method, the PSO algorithm, the entropy weight method, and the variance method, and the validation set data is used in the combined model. Choose the best weights, that is, Prophet takes 0.65745 and LSTM takes 0.34255. Finally, the combined model is compared with the traditional time series model ARIMA model, Prophet model, and LSTM model. The experimental results show that the combined model has better performance and higher accuracy through the combination of the PSO algorithm.

References

[1] Dara S, Tumma P. Feature extraction by using deep learning: A survey[C]//2018 Second International Conference on Electronics, Communication and Aerospace Technology. IEEE, 2018: 1795-1801.
[2] Lv Z, Li J, Dong C, et al. DeepSTF: A Deep Spatial–Temporal Forecast Model of Taxi Flow[J]. The Computer Journal, 2021: 1-16.
[3] Taylor S J, Letham B. Forecasting at scale[J]. The American Statistician, 2018, 72(1): 37-45.
[4] Na G, Lianying S, Xiaoda S, et al. Research on Sales Prediction of Prophet-LSTM Combination Model[J]. Computer Science, 2019, 46(B06): 446-451.
[5] Li Y, Ma Z, Pan Z, et al. Prophet model and Gaussian process regression based user traffic prediction in wireless networks[J]. Science China Information Sciences, 2020, 63(4): 1-8.
[6] Samal K K R, Babu K S, Das S K, et al. Time series based air pollution forecasting using SARIMA and prophet model[C]//2019 international conference on information technology and computer communications. 2019: 80-85.
[7] Wu Q. The forecasting model based on wavelet v-support vector machine[J]. Expert Systems with Applications, 2009, 36(4): 7604-7610.
[8] Du N, Liang X. Short-term water level prediction of Hongze Lake by Prophet-LSTM combined model based on LAE[C]//2021 7th International Conference on Hydraulic and Civil Engineering & Smart Water Conservancy and Intelligent Disaster Reduction Forum. IEEE, 2021: 255-259.
[9] Lv Z, Li J, Dong C, et al. Deep learning in the COVID-19 epidemic: A deep model for urban traffic revitalization index[J]. Data & Knowledge Engineering, 2021, 135: 1-17.
[10] Taylor S J, Letham B.Forecasting at scale[EB/OL].https:/ldoi.org/10.1080/00031305.2017.1380080,2017.
[11] Zhao N, Ying L, Vanos J K, et al. Day-of-week and seasonal patterns of PM 2.5 concentrations over the United States: Time-series analyses using the Prophet procedure[J]. Atmospheric environment, 2018, 192: 116-127.
[12] Beibei C. Research and Application of Combination Prediction Model[D]. Shan Dong University, 2017.
[13] Makridakis S, Spiliotis E, Assimakopoulos V. The M4 Competition: Results, findings, conclusionand way forward[J]. International Journal of Forecasting, 2018, 34(4): 802-808.
[14] Lv Z, Li J, Li H, et al. Blind travel prediction based on obstacle avoidance in indoor scene[J]. Wireless Communications and Mobile Computing, 2021, 2021: 1-9.