

# Research on Ultra-Wideband (UWB) Accurate Positioning under Signal Interference Based on Deep Learning

Ji Zhang, Chenghan Li

North China Electric Power University, Baoding, Hebei, China

---

**Abstract:** UWB (Ultra-Wideband) technology is also called "Ultra-Wideband", also known as pulse radio technology. UWB-based positioning technology has real-time indoor and outdoor accurate tracking capabilities, with high positioning accuracy, which can achieve centimeter-level or even millimeter-level positioning. Based on the provided anchor point ranging information, this paper analyzes normal and abnormal data and establishes an accurate positioning model. For task 1, preprocess the data, export the file, delete invalid data, and fill in missing values. For task 2, it is required to establish models for normal and abnormal data respectively. For task 3, it is consistent with the model obtained in task 2. The only difference is that the coordinates of the four anchor points used for the test data have changed. When the target coordinates are calculated, the anchor point coordinates can be replaced to obtain the model required by task 3. For task 4, use the processed data in task one to establish a mathematical model, and train the model through the integrated learning method to determine whether the collected signal is interference. For task 5, first use the integrated learning model in task four to eliminate the interference data in the data, import the eliminated data into the positioning model of task two for positioning, and add Kalman based on interference recognition to the static estimation algorithm.

**Keywords:** Data processing, ensemble learning, Kalman filter processing, Least squares LS positioning algorithm.

---

## 1. Introduction

With the development of the city, the time people spend indoors for work and activities has gradually increased. For example: quickly find a destination in a large shopping mall or office building; real-time accurate positioning of the location of the elderly and children facilitates care and prevention of loss; positioning the specific locations of various departments, doctors and nurses in the hospital, improving work efficiency and reducing time wasted on the road; At airports, stations and other places with a large flow of people, the flow of people can be monitored more effectively, and work such as regulation and rescue can be carried out in a timely manner [1]. But compared to the empty outdoor environment, the indoor environment is much more complicated. The first is that the indoor space structure is more complex. Multipath effects, NLOS, and the flow of people all greatly increase the difficulty of indoor positioning; secondly, the occlusion of buildings, walls, and furniture greatly weakens the strength of the satellite signal arriving indoors, resulting in positioning. As a result, the error is relatively large, so the satellite positioning system is not suitable for indoor positioning.

Existing indoor positioning can be divided into several categories due to the difference in positioning medium, mainly including infrared indoor positioning based on optics, indoor positioning in visible light; ultrasonic positioning based on sound waves; Bluetooth indoor positioning based on radio wave communication technology, radio frequency identification indoor Positioning etc. Each of the above methods has its own shortcomings, such as limited application scenarios, insufficient positioning accuracy, high power consumption per unit time, and so on. And Ultra Wide Band (UWB) technology [2], as an emerging radio communication technology, has a broad space for

development, and is listed as one of the top ten communication technologies in the future. Aside from the traditional carrier modulation and demodulation method, UWB indoor positioning uses nanosecond-level narrow pulses to transmit signals. At the same time, UWB has better time resolution and strong anti-interference ability, especially its ability to resist multipath interference.

In summary, UWB-based positioning technology has many advantages such as power consumption per unit time, strong noise resistance, high positioning accuracy, etc., which brings multiple advantages when applied to indoor positioning. Nowadays, the demand for indoor positioning is increasing. The more urgent, indoor positioning research based on UWB technology also has great development prospects and application value.

To solve the problem of ultra-wideband (UWB) precise positioning under signal interference[3], using UWB positioning technology (TOF) to collect the distance between the 4 anchor points (anchor) and the target point (Tag), hope to use mathematical modeling (or algorithm) method, regardless of whether the signal interferes, Can give the precise positioning (3D coordinates) of the target (target point). In the experiment, the UWB data of Tag in 324 different positions were collected in experiment scenario 1, under the signal interference-free and signal interference, that is, each position was tested (collected) twice, one signal had no interference, and the other signal had interference (There is shelter between the anchor point and the target point).

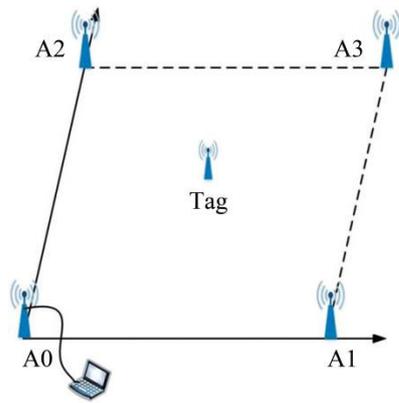


Figure 1. Schematic diagram of the actual measurement environment

## 2. Method

### 2.1. Data Preprocessing

First, import the xlwt library of python to remove the blank lines in the txt file and convert it into an excel file. The normal

data and abnormal data are converted separately. The content of the converted file is shown in the following figure:

Performing data analysis on the converted files, it can be found that the first row of data in each file is invalid data and needs to be removed. Then the overall analysis of the data, it is found that there are missing values in the data, and the missing values are filled. Since this topic is a classification prediction problem, as missing values cannot be simply filled by means or similar values, this article uses high-dimensional mapping methods to fill in, maps attributes to high-dimensional spaces, and uses one-hot coding (one-hot) technology. The attribute value including K discrete value ranges is expanded to K+1 attribute values. If the attribute value is missing, the K+1th attribute value after expansion is set to 1. After comparison, this approach is the most accurate approach. It retains all the information and does not add any additional information. It can completely retain all the information of the original data without considering missing values. The comparison of data before and after processing is shown in the figure below:

A	B	C	D	E	F	G	H	I	J	K	L	M	N
Tag_ID	time	D0	D1	D2	D3	number	number-2	Tag_X	Tag_Y	Tag_Z	D0 <sup>2</sup>	D1 <sup>2</sup>	D2 <sup>2</sup>
0	09053108	760	4550	4550	6300	229	3301	500	500	880	577600	20702500	20702500
0	09053129	760	4550	4550	6300	230	3302	500	500	880	577600	20702500	20702500
0	09053151	770	4550	4550	6300	231	3303	500	500	880	592900	20702500	20702500
0	09053171	780	4550	4550	6300	232	3304	500	500	880	608400	20702500	20702500
0	09053192	780	4550	4550	6300	233	3305	500	500	880	608400	20702500	20702500
0	09053212	770	4550	4550	6300	234	3306	500	500	880	592900	20702500	20702500
0	09053233	770	4550	4550	6300	235	3307	500	500	880	592900	20702500	20702500
0	09053254	770	4550	4550	6310	236	3308	500	500	880	592900	20702500	20702500
0	09053275	770	4550	4550	6300	237	3309	500	500	880	592900	20702500	20702500
0	09053294	770	4550	4550	6300	238	3310	500	500	880	592900	20702500	20702500
0	09053315	780	4550	4560	6310	239	3311	500	500	880	608400	20702500	20793600
0	09053335	770	4550	4560	6310	240	3312	500	500	880	592900	20702500	20793600
0	09053356	770	4550	4550	6300	241	3313	500	500	880	592900	20702500	20702500
0	09053377	770	4550	4550	6300	242	3314	500	500	880	592900	20702500	20702500
0	09053398	770	4550	4560	6300	243	3315	500	500	880	592900	20702500	20793600
0	09053419	780	4550	4560	6300	244	3316	500	500	880	608400	20702500	20793600
0	09053439	770	4550	4570	6300	245	3317	500	500	880	592900	20702500	20884900
0	09053460	760	4550	4560	6300	246	3318	500	500	880	577600	20702500	20793600
0	09053481	760	4550	4560	6300	247	3319	500	500	880	577600	20702500	20793600
0	09053502	770	4550	4550	6300	248	3320	500	500	880	592900	20702500	20702500
0	09053523	770	4550	4550	6300	249	3321	500	500	880	592900	20702500	20702500
0	09053543	760	4550	4550	6300	250	3322	500	500	880	577600	20702500	20702500
0	09053564	760	4550	4550	6300	251	3323	500	500	880	577600	20702500	20702500
0	09053585	760	4550	4550	6300	252	3324	500	500	880	577600	20702500	20702500

Figure 2. Data display after conversion

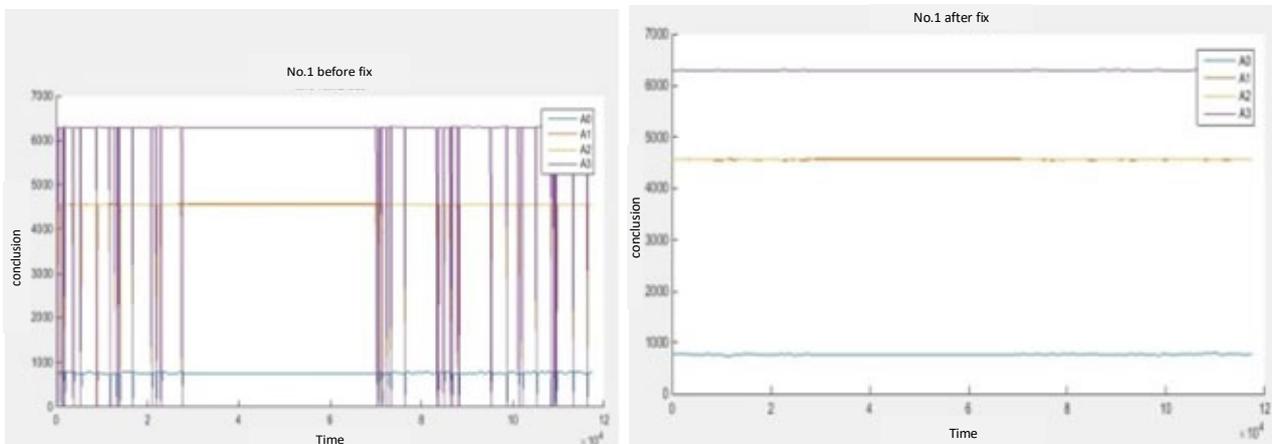
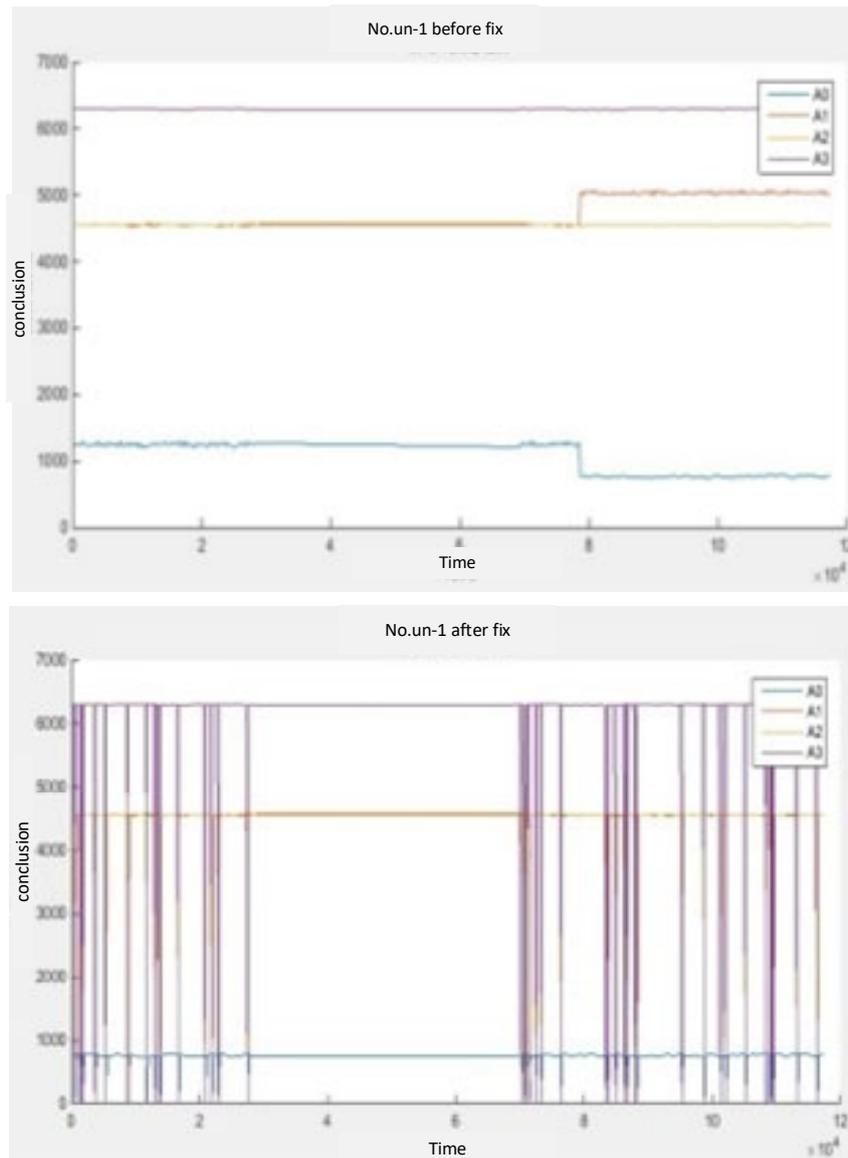


Figure 3. Comparison before and after normal data processing



**Figure 4.** Comparison of abnormal data before and after processing

## 2.2. Solve the Actual Position of the Tag

Task 2 needs to establish a suitable mathematical model (or) algorithm based on the data cleaned in task 1 to solve the actual position of the tag[4]. The data processed by task 1 has been divided into normal data and abnormal data. We need to establish appropriate models to solve the position of the two kinds of data.

For normal data, we only need to take out the mode value of the data to get the position information of the four observation points at the tag. For abnormal data, we must first eliminate interference from it, and then select the data that is tested to be normal before it can be further processed like normal data.

For normal data, extract the distance value of the distance Tag measured by one observation point from the file, calculate the average value of all the distance values of the point, and calculate the measured value for each point in turn. The average value of the distance value of Tag. Then, for each group of data, the difference between the measured distance and the average value of the 4 observation points is obtained in turn, and the 4 differences are summed, and the 10 points with the smallest sum of differences are selected as the distance values sent to the positioning algorithm.

For abnormal data, because there is a ranging error caused

by obstructions, first calculate the average value of each observation point similar to normal data, because the distance measurement error caused by obstructions will only make the distance measurement larger. The average value is compared, and the point where the difference is greater than the threshold value is regarded as an error and marked as abnormal. From the rest of the test, the difference between the measured distance and the average value of the 4 observation points is calculated for each group of data, and the 4 differences are summed, and the 10 points with the smallest sum of differences are selected as the input positioning. The distance value of the algorithm.

LS (Least Squares), also known as Least Squares, is an optimization technique. Take the positioning method based on the time difference of arrival as an example. The essence of this positioning method is the measured TODA value is made hyperbola and solved, the solved equation system is nonlinear [5]. You can convert it to linear first, and then solve it. When the number of base stations of the TDOA positioning method is 4, suppose the coordinates of  $BS_i$  are  $(x_i, y_i, z_i)$ , the coordinates of the target to be measured are  $(x, y, z)$ , and the distance between the target to be measured and the base station is  $r_i$ .

$$\begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} = \begin{bmatrix} \sqrt{(x-x_1)^2 + (y-y_1)^2 + (z-z_1)^2} \\ \sqrt{(x-x_1)^2 + (y-y_1)^2 + (z-z_1)^2} \\ \sqrt{(x-x_2)^2 + (y-y_2)^2 + (z-z_2)^2} \\ \sqrt{(x-x_3)^2 + (y-y_3)^2 + (z-z_3)^2} \end{bmatrix}$$

And:

$$A = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 & r_2 - r_1 \\ x_3 - x_1 & y_3 - y_1 & r_3 - r_1 \\ x_4 - x_1 & y_4 - y_1 & r_4 - r_1 \end{bmatrix}, b = \frac{1}{2} \begin{bmatrix} r_2^2 - r_1^2 - x_2^2 + y_2^2 + x_1^2 + y_1^2 \\ r_3^2 - r_1^2 - x_3^2 + y_3^2 + x_1^2 + y_1^2 \\ r_4^2 - r_1^2 - x_4^2 + y_4^2 + x_1^2 + y_1^2 \end{bmatrix}$$

We can see:

$$\hat{x} = (A^T A)^{-1} A^T b$$

After sorting, the optimal estimated solution optimized by the least square method can be obtained. This optimization method has a small amount of calculation and is widely used. In order to improve positioning accuracy and increase the utilization of redundant positioning information, the LS positioning algorithm can be weighted, that is, WLS (Weighted Least Squares), that is, the weighted least squares method. The principle and process of WLS and LS positioning algorithms are basically the same, the difference is that WLS adds a weighting matrix. Let the weighting matrix be  $W$ , then the solution formula of the weighted least squares method can be sorted as:

$$\hat{x} = (A^T W A)^{-1} A^T W b$$

In order to evaluate and compare the positioning performance of different positioning methods and algorithms, we usually use some quantitative indicators, the most important of which is the positioning accuracy. The positioning accuracy reflects the degree of approximation between the measured value obtained by the positioning algorithm and the actual. Evaluation indicators include root mean square error (RMSE), mean square error (MSE), and so on.

Mean Squared Error, Root Mean Square Error: Mean Squared Error (MSE) refers to the mathematical expectation of the square of the difference between the calculated value and the true value. It reflects the degree of change in the positioning difference. RMSE), also known as standard error, is the most commonly used standard to measure positioning accuracy. It reflects the deviation between the calculated value and the true value, and the value is equal to the arithmetic square root of the mean square error value. The expressions in the two- and three-dimensional cases of MSE and RMSE are as follows:

In a two-dimensional plane:

$$MSE = E[(x-x_0)^2 + (y-y_0)^2]$$

$$RMSE = \sqrt{E[(x-x_0)^2 + (y-y_0)^2]}$$

### 2.3. Determine Signal Interference

After completing task three, it can be found that the model of task three was established under the condition that the signal interference is known, but under normal circumstances, when UWB collects data, the program itself does not know whether the signal is interference. Therefore, it is very important to design a scientific method to successfully determine whether the signal has interference. This is also the focus and difficulty of UWB positioning[6].

The measures taken in this paper are to use the processed data in task one to establish a mathematical model, and to train the model through the integrated learning method to determine whether the collected signals are interfering.

In terms of model construction, the main method used is integrated learning. By analyzing the competition questions, it can be known that this question is to judge whether the signal has interference, so it should be attributed to the classification problem. Set the class feature to represent classification, class = 0 represents no interference, class = 1 represents the signal is interfered and add this feature to the data set.

The main approach is to first introduce algorithms such as logistic regression, support vector machines, K nearest neighbors, decision trees, random forests, integrated learning bagging, stochastic gradients, integrated learning Gradients, and XGB extreme gradient boosting trees, and use the updated data set for gridding Search to obtain the optimal hyperparameters, and use the new data set for cross-validation and training, calculate the AUC score from the training and validation results, combine the three models with the best score results for integrated learning and draw the ROC curve. The flow chart is as follows:

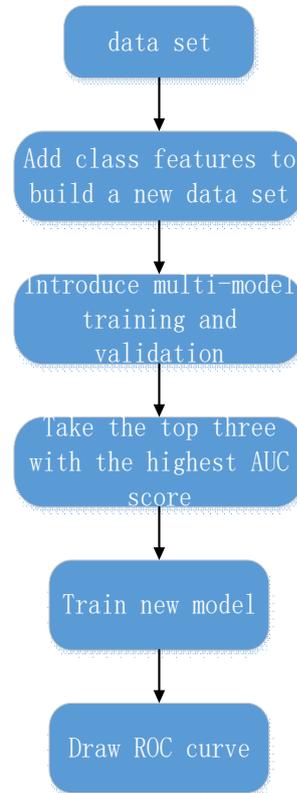


Figure 5. Flowchart display of task four

Logistic Regression (Logistic Regression) is used to deal with the regression problem of the dependent variable as a categorical variable. The common one is the binary or binomial distribution problem, and it can also handle the multi-category problem. The graph of the relationship between the probability and the independent variable of the two classification problem is often an S-shaped curve, as shown in the figure, it is realized by the Sigmoid function.

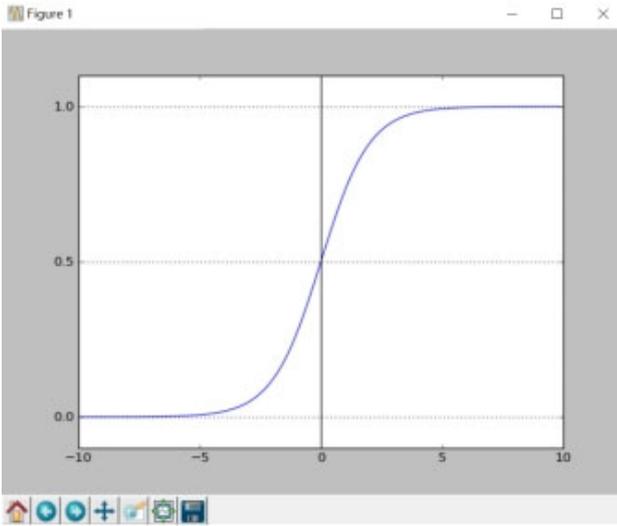


Figure 6. Sigmoid function implementation display

Support vector machine SVM is an original (non-combined) classification algorithm with obvious intuitive geometric meaning and high accuracy. Generally speaking, it is a two-class classification model. Its basic model is defined as the linear classifier with the largest interval in the feature space. That is, the learning strategy of the support vector machine is to maximize the interval, which can finally be transformed into a convex quadratic Solving planning problems. The idea is to convert the problem into a convex optimization problem, which can be simplified by the Lagrange multiplier method, and then solved by the existing algorithm. In complex cases, linearity is inseparable. Use the mapping function to project samples into high-dimensional space to make them linearly separable. Use the kernel function to reduce the amount of high-dimensional calculations.

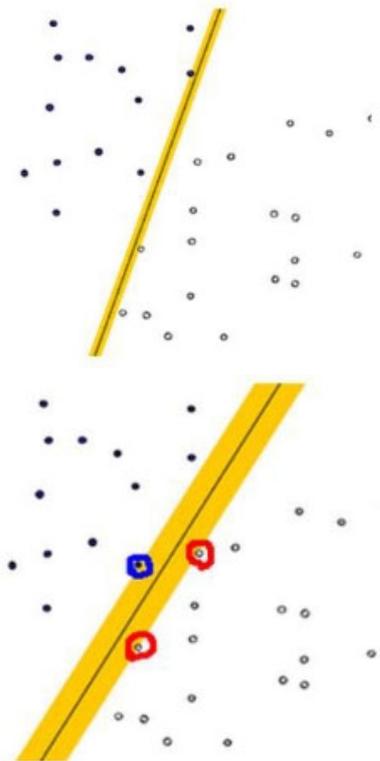


Figure 7. The largest edge hyperplane

k-nearest neighbor (k-NN) is a basic classification and regression method. When classifying, for a new instance, according to the category of the k closest training instances to it, through majority voting, etc. predict. For a given training set, when the k value, distance measurement and classification decision rules (collectively referred to as the three elements) are determined, the model based on the k-nearest neighbor method has been determined. Therefore, it actually uses the training set to divide the feature vector space and does not show the learning process.

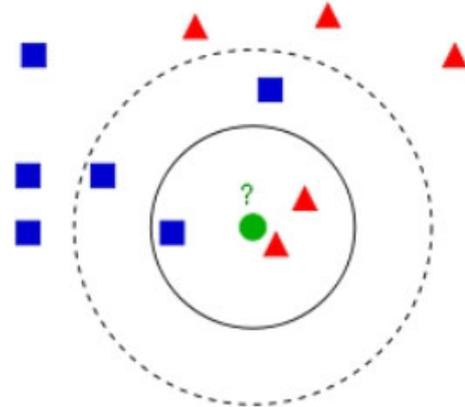


Figure 8. K-value selection

The information gain in the decision tree is defined as follows: Given a sample set D, the entropy of the sample set D before division is certain, represented by H0; Use a certain feature A to divide the data set D, and calculate the entropy of the divided data subset, Expressed by H1, then: information gain = H0-H1, it can also be expressed as:  $g(D,A) = H(D) - H(D|A) = I(D,A)$ . The limitation of information gain: the information gain is biased towards features with more values. Reason: When the value of the feature is large, it is easier to obtain a subset with higher purity according to the feature division, so the entropy after the division is lower, because the entropy before the division is constant, so the information gain is greater, so the information gain Features with more values are more biased. The algorithm that uses information gain as a selection index to generate a decision tree is called the ID3 algorithm.

In XGB, in addition to the tree model, a linear model, such as linear regression, can also be used for integration. There is a loss function in the gradient boosting algorithm. Different from the fixed loss function writing in algorithms such as logistic regression and SVM, the loss function in the integrated algorithm is optional, and the loss function to be selected depends on the problem to be solved. Since this task is to perform classification prediction, you can choose error rate error or log loss  $\log\_loss$ . As long as the selected function is a differentiable function that can represent a certain loss, it can be the loss function in XGB. The objective function of XGB is written as: traditional loss function + model complexity:

$$Obj = \sum_{i=1}^m l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k)$$

$$\hat{y}^{(t)} = \sum_k f_k(x_i) + f_t(x_i)$$

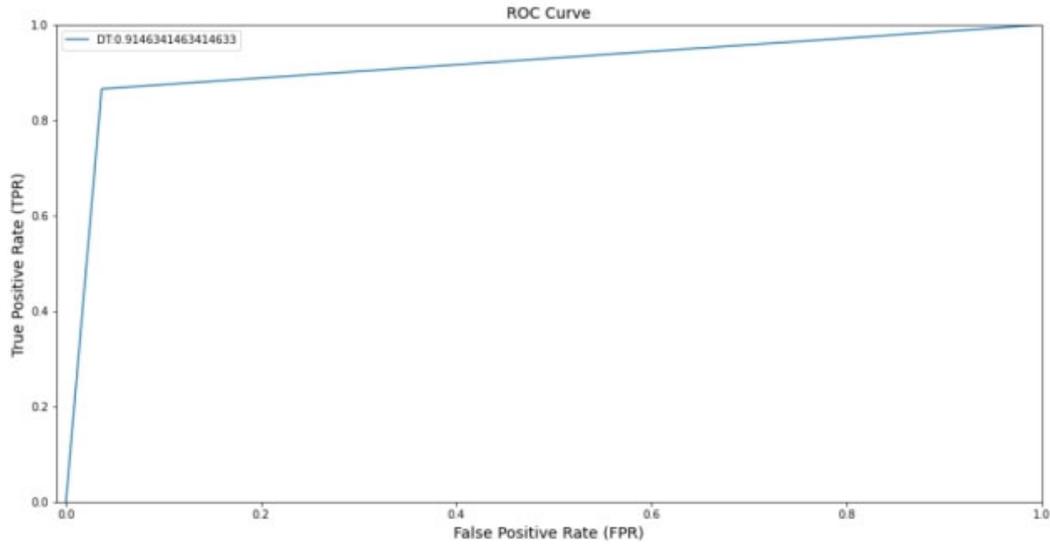


Figure 9. ROC curve of decision tree model

Perform preliminary training through the above models to obtain their respective hyperparameters, put all the most hyperparameter models together, call accuracy, F1-score and

other functions to calculate the values of various indicators, and put each value into the DataFrame. The indicator values are shown in the figure below:

	Accuracy	F1-score	Recall	Precision	AUC_ROC
LR	0.944162	[0.9473684210526315, 0.9405405405405406]	[0.9705882352941176, 0.9157894736842105]	[0.9252336448598131, 0.9666666666666667]	0.943189
KNN	0.928934	[0.9339622641509434, 0.9230769230769231]	[0.9705882352941176, 0.8842105263157894]	[0.9, 0.9655172413793104]	0.927399
SVC	0.944162	[0.9473684210526315, 0.9405405405405406]	[0.9705882352941176, 0.9157894736842105]	[0.9252336448598131, 0.9666666666666667]	0.943189
DT	0.908629	[0.9150943396226415, 0.901098901098901]	[0.9509803921568627, 0.8631578947368421]	[0.8818181818181818, 0.9425287356321839]	0.907069
RFC	0.923858	[0.9302325581395348, 0.9162011173184357]	[0.9803921568627451, 0.8631578947368421]	[0.8849557522123894, 0.9761904761904762]	0.921775
Bagging	0.928934	[0.9345794392523364, 0.9222222222222223]	[0.9803921568627451, 0.8736842105263158]	[0.8928571428571429, 0.9764705882352941]	0.927038
SGD	0.923858	[0.9275362318840579, 0.9197860962566845]	[0.9411764705882353, 0.9052631578947369]	[0.9142857142857143, 0.9347826086956522]	0.923220
XGB	0.939086	[0.9433962264150944, 0.9340659340659342]	[0.9803921568627451, 0.8947368421052632]	[0.9090909090909091, 0.9770114942528736]	0.937564

Figure 10. Display of various indicators

## 2.4. Determine Signal Interference

The idea of the topic is to locate the motion trajectory. According to the requirements of the topic, the data that can be referred to are the positioning model of the static point and the motion law of the target point itself. At the same time, the data in Annex 5 has interference during the collection process, and it is necessary to eliminate the influence of abnormal data on the normal experimental results, and finally output the motion trajectory diagram. In this task, the title has clearly informed that random interference will occur, which means that there are both normal data and abnormal data in the timeline. Therefore, you can use the integrated learning model in task four to filter the interference data, and then use the task The second model constructs a positioning model based on interference data to determine its coordinates, and combines the coordinates of normal data to fit the final motion trajectory diagram.

In terms of model establishment, this article first uses the integrated learning model in task four to eliminate

interference data in the data, and imports the eliminated data into the positioning model of task two for positioning. At the same time, it adds interference recognition based on the static estimation algorithm. Kalman filter processing.

Simply put, the Kalman filter is an "optical recursive data processing algorithm". The Kalman filter is the best, most efficient and even the most useful for solving most problems. The characteristics of the Kalman filter: (1) The object processed by the Kalman filter is a random signal; (2) The processed signal has no distinction between usefulness and interference. The purpose of filtering is to estimate all processed signals (different from Wiener filtering); (3) The white noise excitation and measurement noise of the system are not objects that need to be filtered out. Their statistical characteristics are the information that needs to be used in the estimation process; (different from the least squares method) (4) The algorithm is Inferred, and use the state space method to design the filter in the time domain, which is suitable for multi-dimensional random processes; (5) The estimated quantity can be either stationary or non-stationary; (6) In the

estimation process, only need to consider The statistical characteristics of the process noise and measurement noise and the current state of the system. (It costs less computer computing resources)

### 3. Experimental Results

Perform preliminary training through the above models to

obtain their respective hyperparameters, put all the most hyperparameter models together, call accuracy, F1-score and other functions to calculate the values of various indicators, and put each value into the DataFrame. The indicator values are shown in the figure below:

	Accuracy	F1-score	Recall	Precision	AUC_ROC
LR	0.944162	[0.9473684210526315, 0.9405405405405406]	[0.9705882352941176, 0.9157894736842105]	[0.9252336448598131, 0.9666666666666667]	0.943189
KNN	0.928934	[0.9339622641509434, 0.9230769230769231]	[0.9705882352941176, 0.8842105263157894]	[0.9, 0.9655172413793104]	0.927399
SVC	0.944162	[0.9473684210526315, 0.9405405405405406]	[0.9705882352941176, 0.9157894736842105]	[0.9252336448598131, 0.9666666666666667]	0.943189
DT	0.908629	[0.9150943396226415, 0.901098901098901]	[0.9509803921568627, 0.8631578947368421]	[0.8818181818181818, 0.9425287356321839]	0.907069
RFC	0.923858	[0.9302325581395348, 0.9162011173184357]	[0.9803921568627451, 0.8631578947368421]	[0.8849557522123894, 0.9761904761904762]	0.921775
Bagging	0.928934	[0.9345794392523364, 0.9222222222222223]	[0.9803921568627451, 0.8736842105263158]	[0.8928571428571429, 0.9764705882352941]	0.927038
SGD	0.923858	[0.9275362318840579, 0.9197860962566845]	[0.9411764705882353, 0.9052631578947369]	[0.9142857142857143, 0.9347826086956522]	0.923220
XGB	0.939086	[0.9433962264150944, 0.9340659340659342]	[0.9803921568627451, 0.8947368421052632]	[0.9090909090909091, 0.9770114942528736]	0.937564

Figure 11. Display of various indicators

The base model is selected through the AUC value, and the visualized results of the AUC value are as follows, and LR, SVC, and XGB are selected as the base model through analysis.

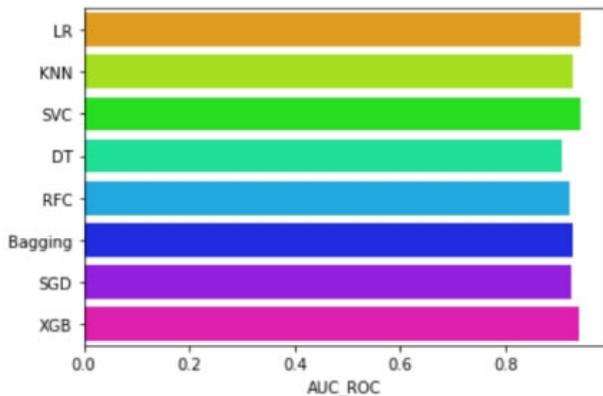


Figure 12. Visualization of AUC value

Integrating the three base models of LR, SVC, and XGB as a model, the prediction results of the new model obtained through training are significantly higher than the individual prediction results of the previous models.

	precision	recall	f1-score	support
0	0.93	0.98	0.95	102
1	0.98	0.92	0.95	95
accuracy			0.95	197
macro avg	0.95	0.95	0.95	197
weighted avg	0.95	0.95	0.95	197

Figure 13. Integrated learning prediction score

Table 1. Integrated learning prediction results

Signal	Location
Normal Signal 1	Estimated location [1550,2576,1253]
Interference signal 2	Estimated location [2062,4453,1662]
Normal Signal 3	Estimated location [4108,2530,1476]
Interference signal 4	Estimated location [3273,3778,1350]
Suspected interference5	Estimated location [1432,2545,1009]
Suspected interference6	Estimated location [2050,4531,1688]
Suspected interference7	Estimated location [3273,3778,1350]
Suspected interference8	Estimated location [3368,3681,640]
Normal Signal 9	Estimated location [4177,2394,826]
Normal Signal 10	Estimated location [2453,1026,2223]

According to the requirements of task five, the data that can be referred to are the positioning model of the static point and the motion law of the target point itself. At the same time, the data in Annex 5 has interference during the collection process, and it is necessary to eliminate the influence of abnormal data on the normal experimental results, and finally output the

motion trajectory diagram. In this task, the title has clearly informed that random interference will occur, which means that there are both normal data and abnormal data in the timeline. Therefore, you can use the integrated learning model in task four to filter the interference data, and then use the task The second model constructs a positioning model based on interference data to determine its coordinates, and combines the coordinates of normal data to fit the final motion trajectory diagram.

The so-called filtering is to extract the desired signal from the mixed signals. Signals can be divided into deterministic signals and random signals. Deterministic signals: can be expressed as a definite time function, and its magnitude (with a definite frequency spectrum) at any moment can be determined. Random signal: It cannot be described by a definite mathematical relationship, cannot predict any instantaneous value, and its change obeys the statistical law (the frequency is uncertain, and the power spectrum is determined).

After the analysis of the algorithm, the trajectory of the movement, the distance measurement result, and the data analysis are output as pictures:

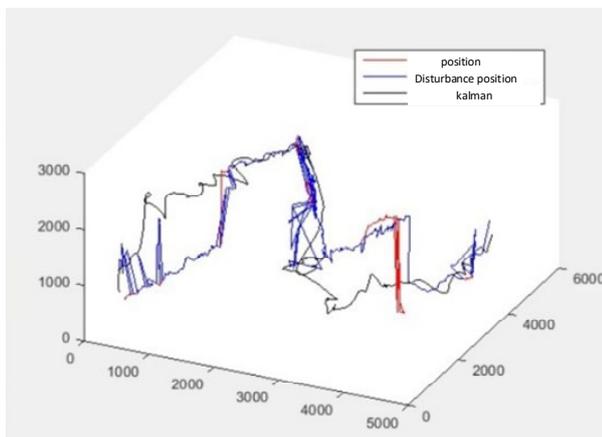


Figure 14. Motion trajectory

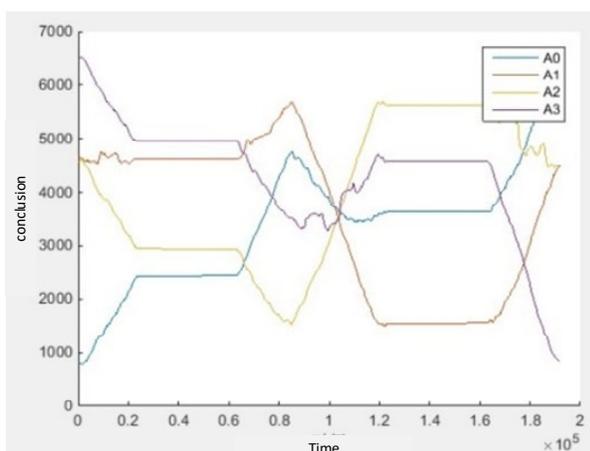


Figure 15. Ranging results

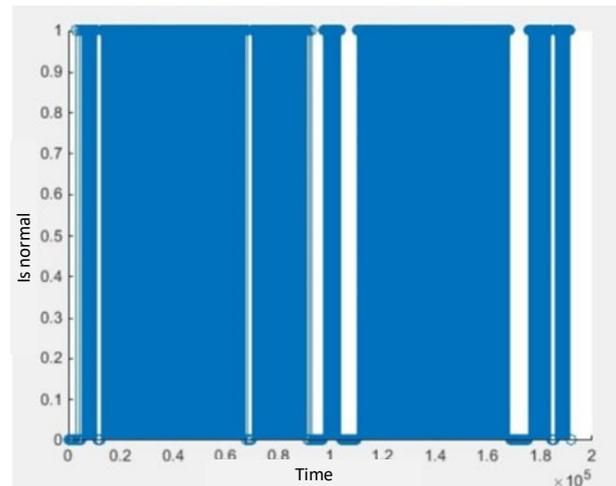


Figure 16. Data analysis

## 4. Conclusion

This article first preprocesses the data, exports the file, deletes invalid data, and fills in the missing values; secondly, it is required to establish models for normal and abnormal data respectively, and select the 10 points with the smallest difference for the smallest improvement based on the TDOA algorithm. The two-fold LS positioning algorithm is used to locate the target; the LR, SVC, and XGB models are trained again through the integrated learning method to determine whether the collected signal is interference. Finally, the integrated learning model is used to eliminate the interference data in the data, and the eliminated data is imported into the positioning model of task two for positioning. At the same time, the Kalman filtering process based on interference recognition is added on the basis of the static estimation algorithm. Through experimental comparison and analysis, the accuracy of the algorithm proposed in this paper is verified, and the integrated learning model model is adopted to improve the accuracy and generalization ability. However, the time required for training is longer and the computational cost is higher. Therefore, reducing training time should be the focus of the next step.

## References

- [1] Zhao Jielei. Research and implementation of ultra-wideband indoor positioning algorithm. 2019. Zhengzhou University of Light Industry, MA thesis.
- [2] Kok M, Hol J D, Schon T B. Indoor Positioning Using Ultrawideband and Inertial Measurements[J]. IEEE Transactions on Vehicular Technology, 2015, 64(4): 1293—1303.
- [3] Arias-De-Reyna E. A Cooperative Localization Algorithm for UWB Indoor Sensor Networks[J]. Wireless Personal Communications, 2013, 72(1):85—99.
- [4] Miao Xiren, Fan Jianwei, etc. Ultra-wideband indoor positioning method based on improved extreme learning machine under abnormal base station conditions, Journal of Sensor Technology, 2020, 33(10): 1-10.
- [5] Su Xiang. Research on hybrid TDOA/AOA indoor positioning technology based on extended Kalman filter[J]. Digital Technology and Application, 2013(8):56-57.
- [6] Yan Jiaqi. Algorithm and error analysis of indoor positioning system based on UWB [D]. Harbin Institute of Technology, 2020.