Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 2458-9403
Vol. 6 Issue 12, December - 2019

Prediction Of Fault Type Of Vacuum Coating Machine Based On Support Vector

Zhong Liangwei, Li Zhenyu, NiJing, ZhuYi
University of Shanghai for Science and Technology
USST
Shanghai, China
zlvcad@126.com

Dai Xiuhai
OPTORUN (SHANGHAI) CO., LTD
Shanghai, China
daixh@optorun.sh.cn

***Abstract*—The working parameters of coating machine have great influence on coating quality, for the real-time monitoring and fault prediction of operating state parameters of coating machine, a fault prediction method of vacuum coating machine based on support vector machine theory is proposed, based on the open source machine learning tool scikit-learn, RBF kernel support vector algorithm was applied to analyze a large number of working data of the existing coating machine, build a monitoring model of the coating machine, and optimize the model. The method of taking the real-time data collection function as the system service is studied, which provides a reliable guarantee for the remote interaction of the web end coating machine.**

Keywords—component; Coating machine; Scikit - learn; Support vector; Failure prediction

I. INTRODUCTION

Vacuum coating machine coating process of the change of parameters are many, mainly including cavity temperature, coating rate, ion source parameters, pressure size. Coating machine coating quality depends on the user coating set parameters and the stability of the coating machine, if the coating machine parameters change can not be timely detected and adjusted, will produce a large number of defective products. In this paper, massive data of the coating machine in operation are used to build a classifier to predict the fault types of the coating machine with the help of machine learning, It provides a feasible method to improve the quality of coating machine.

At present, machine learning methods mainly include support vector machine, decision tree, k-nearest neighbor, neural network and so on. In this paper, support vector machine method is selected. Based on Python3.5 environment, scikit-learn, an open source machine learning tool, was used to establish the monitoring model of the coating machine, and the running state and fault types of the vacuum coating machine were analyzed. The real-time data collected by The coating machine are processed and stored in The database, and The web end displays The corresponding information through The collection and processing.

II. DATA PROCESSING AND MODEL ESTABLISHMENT

A. Data Processing

In this paper, a large amount of data of vacuum coating machine collected in the early stage was processed into CSV file, and the corresponding fault category was marked as the corresponding fault number, where the number 0 indicated the normal state. The document includes 11 parameters that have a significant impact on the coating machine as characteristics, including: monitor piece temperature(mt), penning vacuum gauge pressure(ppg), ionization vacuum gauge FIL pressure(pigf), Beam current(bc), Beam voltage(bv), rf forward power 1(rfp1), emission current(rec), rf forward power 2(rfp2), gas 1 flow(g1), gas 2 flow(g2), gas 3 flow(g3). The processed CSV file is used as the training file for the model. The Python program extracts the training files for the model from the processed CSV file by introducing pandas. Because of the support vector machine (SVM) method was highly sensitive to the scale of data, the parameters of the coating machine characteristics between the data is not in an order of magnitude, so the scikit - learn provided in MinMaxScaler to preprocess the data, make all features just lies between 0 and 1, so as to improve the accuracy of monitoring model, reducing the number of memory and time overhead[1,2].

B. Model Establishment and Tuning

By introducing train_test_split in sicikit-learn, parameters of the coating machine can be divided into two parts: one is used to build the training set of the supervision model of the coating machine, and the other is used to evaluate the performance of the supervision model of the coating machine. Data from the training set was used for model training. Machine learning models in scikit-learn were implemented in their respective classes, which were called Estimator classes. SVM classification algorithm was implemented in SVM module SVC class. To train classifiers in higher dimensional space and reduce computational overhead, use kernel trick. For support vector machines, there are generally two methods to map data to higher dimensional space: one is polynomial kernel; The other is radial basis function (RBF) [3,4]. In this paper, RBF SVM method is used to predict the fault type of coating machine, which includes two very important parameters: regularization parameter C and core width gamma, which are of great significance to

the generalization ability of the monitoring model of coating machine to be constructed. The regularization parameter C and core width gamma were adjusted to improve the generalization performance of the monitor model of the coater, and the combination of all parameters was tried through cross-validation grid search, the principle of which is shown in figure 1.

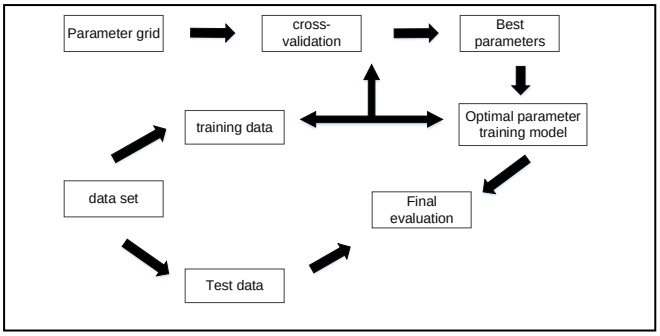

Fig. 1. *Cross-validation.*

Scikit - learn GridSearchCV class, in the form of the estimator is implemented cross-validation grid search, choose to use the model of support vector classification, set up the parameter search grid, select effective cross validation strategy. After GridSearchCV class was instantiated, fit method was called to conduct cross-validation grid search, and the monitor model of coating machine was fitted on the whole training set by the best combination of regularization parameter C and core width gamma.

Due to the difference in the order of magnitude between the characteristics of different parameters of coating machine, the parameters of coating machine should be scaled. If the model is built directly on the scaled training set through cross-validation grid search, the problem of using this model for the test set is not ideal. This is because in each partition of cross-validation, the scaled training set is divided into the training part and the test part, the test part tests the performance of the model constructed by the training part, and the test part is part of the scaled training set. In general, part of the information has been leaked to the modeling process [5]. In this paper, the processing steps MinMaxScaler and SVC are merged into a single scikit-learn estimator by using make_pipeline. After the instantiation of the constructed make_pipeline, it is combined with GridSearchCV. The value of regularization parameter C of SVC is set as 0.1, 1, 10, 100,1000, and the value of gamma of core width is 0.001, 0.01, 0.1, 1, 10, 100.The cross-validation grid search was carried out with the strategy of 10-fold layered cross-validation, and the specific steps to establish the supervision model of the coating machine were shown in figure 2 [6,7,8].

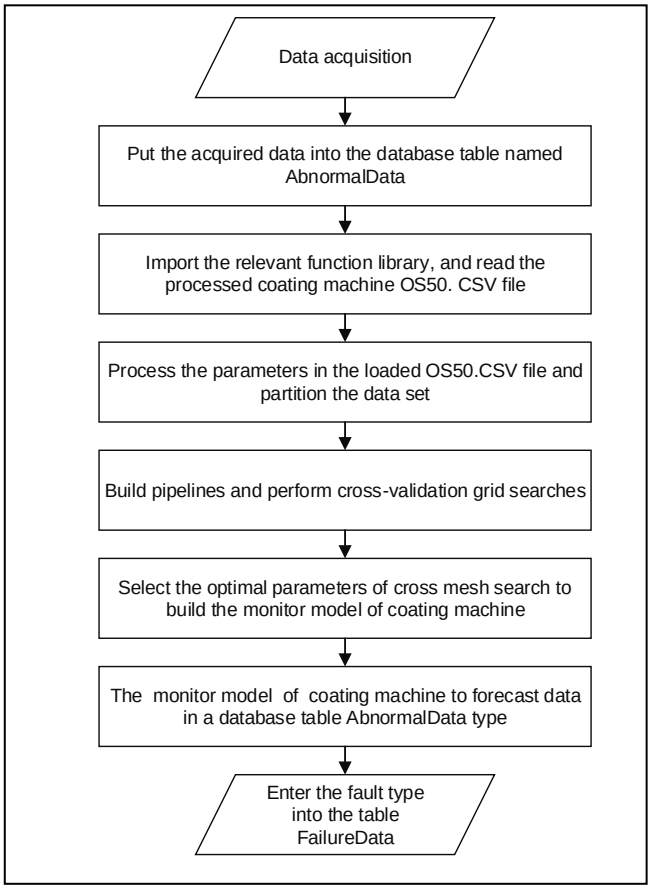

Fig. 2. *Process of constructing the supervision model of coating machine*

After cross validation grid search, realize the visualization of heat map. The dark area in figure 3 represents the poor performance state of the model, while the light area represents the good performance state of the model. The parameters on the grid represent the accuracy of the model when the corresponding regularization parameter C and the kernel width gamma are selected. It can be seen that the highest accuracy is 99.2% when the regularization parameter C is 100 core width gamma 1, so this parameter is selected to build the monitoring model of the coating machine. The test set is applied to the built supervision model of the coating machine to test its robustness, and the test accuracy is 97.8%..

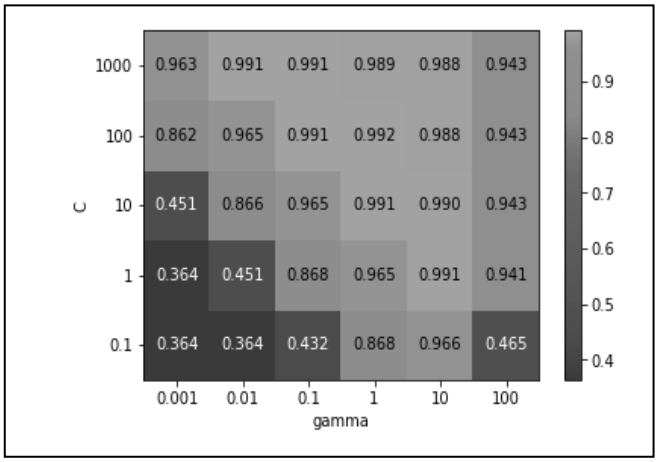

Fig. 3. *Cross - validation accuracy of the heat map*

C. Evaluation and Analysis of Experiments

The data set of the coating machine is an unbalanced data set. In many moments, the coating machine is in normal state, and only in a few abnormal moments can the corresponding fault type occur. Therefore, it is not possible to simply evaluate the model by matching the percentage of predicted value and real value. For such unbalanced data sets, the accuracy of multiple classification is defined as the proportion of correctly classified samples. In this paper, classification_report function is used to calculate the precision, recall rate and f-score of each category through classification report [9,10,11], as shown in table 1:

TABLE I. CLASSIFICATION REPORT OF SUPERVISION MODEL OF COATING MACHINE

	precision	***recall***	***f1-score***	***support***
0	0.98	0.96	0.97	138
1	1.00	1.00	1.00	24
2	0.96	1.00	0.98	26
3	0.95	1.00	0.98	21
4	1.00	1.00	1.00	24
5	1.00	1.00	1.00	25
6	1.00	1.00	1.00	24
7	1.00	1.00	1.00	19
8	0.96	1.00	0.98	27
9	0.94	0.92	0.93	37
micro avg	0.98	0.98	0.98	365
macro avg	0.98	0.98	0.98	365
weighted avg	0.98	0.98	0.98	365

As can be seen from the classification report in table 1, the precision of the correct type 0 in the test set is 98% and the recall rate is 96%.;The precision and recall rate of error types 1, 4, 5, 6, 7 are all good, reaching 1. There is no confusion in these categories; The recall rate of error type 2 and 3 reached 1. There were no false counterexamples in these two types. According to the harmonic average f- score of precision and recall rate, the supervision model of coating machine has met the requirements.

III. APPLICATION OF THE MODEL

The real-time data of coating machine is transmitted to the SQLServer by Wincc in real time. Web side will be used in real-time data processing, processing of data into a database table called AbnormalData table. Pymssql module is introduced into the python script file to connect with SQLServer. After the connection, the use of python for used to send instructions, by iterating through the database table AbnormalData, will AbnormalData the wrong data to the last column in the table type identification for 999 (meaning is: to determine the fault data) to extract the data. By using the method in numpy, the extracted data are merged into vectors that can be recognized by the monitor model of the coating machine. The processed vector is passed into the function constructed by predict method of the monitor model of coating machine to judge it. When the judgment result is normal, delete this line of data in the table AbnormalData directly, if the judgment for the wrong type, the line of your table AbnormalData data into a database table called FailureData table, and in the last column in table FailureData marked the wrong type, finally will be delete the corresponding data in table AbnormalData. The Web side can alarm and display the fault and fault type by processing the data in the table FailureData.

Connect the coating machine for testing, coating machine monitoring model based on table AbnormalData testing, and save the results in table FailureData. Where the number is 2,..., n-1 and n are fault data, and the fault types are respectively fault types 5,..., fault type 3, fault type 1,tables AbnormalData and FailureData part of the test information as shown in table 2 and table 3.

TABLE II. THE DATABASE TABLE ABNORMALDATA

Time	***bv***	...	***rec***	***Type***
1	1300	...	1951	999
2	1287	...	1902	999
...	...	...	...	...
n-1	1302	...	1951	999
n	1301	...	1949	999

TABLE III. DATABASE TABLE FAILUREDATA

Time	***bv***	...	***rec***	***Type***
2	1287	...	1902	5
...	...	...	...	...
n-1	1302	...	1951	3
n	1301	...	1949	1

IV. PACKAGE AND REGISTER

The established coating machine supervision model and the operation of the database is implemented on the server, Python is a scripting language, need an interpreter to interpret before the execution. Use pyinsatller to package the scripts as exe executable files , and through NSSM software will coating machine monitoring exe executable file registration as system services, the start type is set to automatic, can be realized in the state of server boot running coating machine monitoring program, so as to safely deploy coating machine monitoring project.

V. CONCLUSION

Machine learning technology has been applied in many fields. This paper tries to apply machine learning to the fault prediction of coating machine. By using RBF kernel support vector machine (SVM) algorithm,

the supervision model of coating machine is established and optimized. The real time data of the coating machine collected from the server is processed effectively through the constructed model, which provides a reliable guarantee for the interaction of the web end. At the same time, the script can be registered as the system service on the server, to achieve the function, but also to ensure the stability of the operation, the development of coating machine has practical significance.

Acknowledgment

This work was supported by Special Fund for Science and Technology Innovation of Shanghai Baoshan Science and Technology Commission (Grant Nos. 17-c-17).

REFERENCES

[1] Chen xiaodong, wang lei, he he, song jian. Control system principle and circuit fault analysis method of coating machine [J]. Power technology,2011,35(09):1169-1170.

[2] Cao wenpeng, zhou zhenggan, feng zhanying. Application of PLC in multi-arc ion coating machine system [J]. Mechanical and electrical engineering technology,2004(12):65-67+83.

[3] G. Varoquaux,L. Buitinck,G. Louppe,O. Grisel,F. Pedregosa,A. Mueller. Scikit-learn[J]. GetMobile: Mobile Computing and Communications,2015,19(1).

[4] Yang fujian. New trend of machine learning in the era of big data [J]. Electronic technology and software engineering,2019(19):155-156.

[5] An Introduction to Machine Learning with Python by Andreas C.Muller and Sarah Guido(O'Reilly).Copyright 2017 Sarah Guido and Andreas Muller,978-1-449-36941-5.

[6] N. S. N. Shaharum,H. Z. M. Shafri,W. A. W. A. K. Ghani,S. Samsatli,B. Yusuf,M. M. A. Al-Habshi,H. M. Prince. IMAGE CLASSIFICATION FOR MAPPING OIL PALM DISTRIBUTION VIA SUPPORT VECTOR MACHINE USING SCIKIT-LEARN MODULE[J]. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,2018,XLII-4/W9.

[7] Hu Lin, liu tingting, li huan, cui yunpeng. Prospect of machine learning and its application in agriculture [J]. Agricultural library information,2019,31(10):12-22.

[8] Intelligent fault diagnosis based on machine learning [J]. Electronic finance,2018(09):85-87.

[9] Yang Yi, Li Jianguo, Ge Fangzhen. Empirical research on spam message filtering method based on scikit learn [J]. Journal of Huaibei Normal University (NATURAL SCIENCE EDITION)c,2016,37(04):39-41.

[10] Miao qiang, jiang jing, zhang heng, luo chong. Aviation intelligent engine under the background of industrial big data: opportunities and challenges [J]. Journal of instrumentation,2019,40(07):1-12.

[11] Lu yanxia. Analysis and application of machine learning algorithm in fan fault prediction [D]. North China electric power university (Beijing),2018.