Research on Fault Diagnosis Method of Aircraft Engine on Basis of Data Classification Prediction

Hao JiangWei[1a*]，Lu Bin[1]

[1]North China Electric Power University, Baoding 071003, China

[a]haojiangwei@126.com

Abstract. In this article, we introduced the classification and prediction algorithm of data mining into aircraft engine fault diagnosis, then analysis and compare the Random Forest algorithm and Neural Network algorithm.The simulation of these algorithms and experimental results show that the two methods have their own advantages and disadvantages. Combining both of them into system that fault types and potential problems of engines can be found more accurately.

1 Introduction

With the rapid development of Internet technology, innovative technology and cutting-edge concepts continue to affect people's lives. The widely used of artificial intelligence technology promote significant changes in the human-computer relationship, and researchers have begun to more in-depth study of intelligent fault diagnosis technology. Civil aviation market has grown, behind the boom in the aviation industry, aviation maintenance and repair also assume an important role. Twenty percent of the cost of flying in airline is repair fees. Aircraft flight safety is the lifeline of the airline, aircraft engine as the "heart" of aircraft is the basis for safe flying. Aircraft engine complex structure working at harsh conditions of high speed and high temperatures prone to all kinds of mechanical failure. Furthermore, as a member of an aircraft，health condition of the engine will have a direct impact on the safety and punctuality of flight.According to statistics, in all kinds of air accidents caused by various factors, the reason for the engine malfunction percentage is generally 25% to 30%，often lead to disastrous accident in flight [1]. Therefore, for aero-engine safe, economical and efficient operation, it is necessary to fully understand the overall health of aircraft engines, and grasp its variation during operation, at the same time implement status monitoring and fault diagnosis to key components of the engine. But the relationship is nonlinear, uncertain and erratic between aircraft engine state parameters and fault type, traditional methods can not meet the requirement of wear fault diagnosis.

Classification and prediction is the main method of data mining technology. It has widespread attention and research that how to use data mining techniques to analysis and learning known data efficiently which can obtain accurate and effective classification to prediction information about future trends, and its achievements are rapidly put into practical application. From applied research of classification and prediction in data mining, there are many algorithms, such as decision trees, Bayesian, k-nearest neighbor classification, Neural Networks, support vector machines. The data mining prediction and classification algorithm is introduced aircraft engine fault diagnosis in this article. According to four engine state parameters of Boeing 747-400(engine model PW-4000). The initial diagnosis: low pressure compressor speed NL; high pressure compressor speed NH; exhaust gas temperature EGT; fuel flow FF [3]. The above status parameter and corresponding the type of fault is introduced to classification and prediction algorithm for simulation experiment.

2 Random Forest

Random Forest algorithm is a new classification and prediction models that Leo Breiman proposed in 2001, It is classifier ensemble algorithm based on the decision tree. Establish a forest with a random manner and integrated classifier algorithm consist of a plurality of decision

tree$^{\{h(x,\theta i),i=1,2,...,k\}}$,then input x,only one vote for every tree to the category that it considers most appropriate.

Each decision tree is not associated with each other in a Random Forest.In the process of the establishment of each decision tree,we need to pay attention to sampling and completely split.First, two random sampling process.Random Forest make the input data to the line and column sampling.For the row sampling, using the way of take out and put it back, there may be duplicate samples in the sample set sampled.Assuming that the number of input samples is N,then the sampled samples is N.In the training allows input samples of each tree are not all of the samples, so that the relative is not prone to over-fitting.Then the column samples,select the m(m << M) from the M feature.After the data is sampled using a full split after ways to build a decision tree, a leaf node of the tree is either unable to continue splitting or all samples inside are pointing to the same category[4].The algorithm has a few parameters need to be adjusted, not have to worry over-fitting,classification fast.And it can have the features that more efficiently handle large data sample, estimate which feature is more important in the classification, strong noise immunity and the better performance of the data set[2].

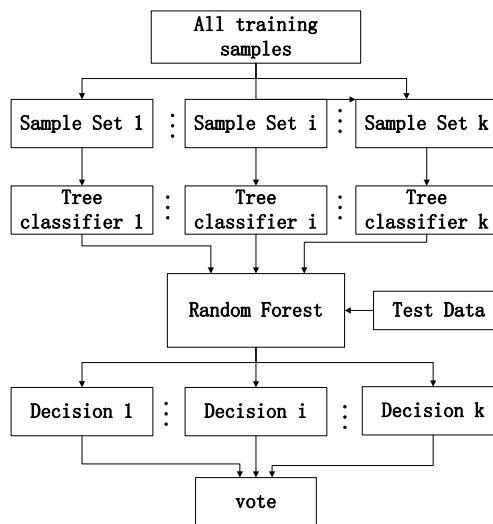Aircraft engine fault diagnosis process based on Random Forest shown in Figure 1.

Figure 1. Random Forest flowchart

3 BP Neural Network

BP Neural Network is a mathematical model that applied similar brain synaptic coupling structure to processing information. It is a multilayer forward network in one-way propagation. It is a Neural Network with three layers or more than three layers, including the input layer, the middle layer (hidden layer) and the output layer. The upper and lower layers achieve full connection, and there is no connection between each layer of neurons. The BP algorithm includes two aspects: the forward propagation of the signal and the back propagation of the error. The direction is from the input to the output direction when calculate the actual output, while the direction is from the output to the input when amending the weights and thresholds.

Aircraft engine fault diagnosis process based on Neural Network is shown in Figure 2.
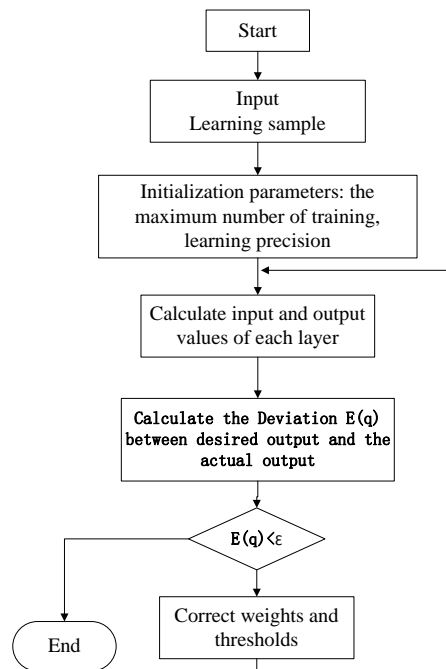
Figure 2. Neural Network flowchart

4 Aircraft engine fault diagnosis simulation

Simulation of this article use programming language Golang that Google developed with the characteristics of compiled, parallelized and garbage collection to achieve Random Forest algorithm and Neural Network algorithm.

4.1 Input Features. According to the experience of aircraft maintenance personnel of engine failure, Boeing 747-400 (engine model PW4000) engine fault can be initial diagnosis according to the following four parameters: low pressure compressor speed NL; high pressure compressor speed NH; exhaust gas temperature EGT; fuel flow FF. Fault diagnosis is use these limited parameters characteristic information provided to determine the state of the engine fault[3].We selected 12 features of literature [3],that is a typical failure mode as the learning sample data, as shown in Table1.

Table 1. PW-4000 typical engine failure

No.	Fault Name	No.	Fault Name
1	3.0 deflate valve malfunction	7	TCC system failure
2	3.5 deflate valve malfunction	8	BETA partial opening angle
3	3.0 and 3.5 deflate valve malfunction	9	High pressure turbine failure
4	Combustion fault	10	The reduce of the low pressure compressor efficiency
5	Eight deflate air leak	11	The reduce of the high pressure compressor efficiency
6	Fifteen deflate air leak	12	A turbine failure

4.2 Learning and test samples. Random Forest algorithm:
On four parameters tested: NL, NH, EGT, FF, In this paper, each fault selected 20 samples, a total of 240 learning samples, make the input and target sample as a algorithm entrance to classify learning.

Take 100 independent tree in the trials to form a Random Forest. According to typical failure modes sample input and output in the literature [3].50 groups chosen as the test sample to the test data input to Random Forest algorithm written in Golang. The results showed that 100 Decision Tree group vote on 50 samples in the correct rate of 96%, training time is 0.59s, as shown in Figure3.

Figure3. Result of the Random Forest algorithm program

In this paper, we have done lots of experiments of Random Forest algorithm in aircraft engine fault diagnosis, when the number of samples is set to select 100 groups, a small number, little number of decision trees or no significant differences in the sample, the test data troubleshooting error rates as high as 30% or more. Therefore, whether the Random Forest accurate decisions or not is determined by the correct learning sample, appropriate number of samples and reduce the relevance of the sample data.

Neural Network algorithm:

On four parameters tested: NL, NH, EGT, FF. In this paper, each fault selected 10 samples, a total of 120 learning samples, make the input and target sample as a algorithm entrance to classify learning.

Corresponding to Table 1 of 12 typical fault mode, perform training with the parameters that input vector dimension 4, output vector dimension of 12 species, the actual use of hidden layer neuron number 10, the learning rate is 0.01, precision control parameter 0.001 and cycles 10,000 times. After training and input the testing data to get the probability type of fault in each data set , the type of fault is determined by the probability of test data.

A set of test results for the training data are shown in Figure4.

Figure 4. The result of eural Network and Random Forest

By contrast experiment multiple sets of data, Neural Network algorithm can basically be able to determine the type of engine failure, you can discover the advantages of potential failures by the probability of each type of fault, but the accuracy is far less than the Random Forest algorithm. Use a combination of both into the system, so that we can find type of engine fault and potential problems more accurate.

4.3 System implementation. Front developing reference MVC mode and used AngularJs framework combined with HTML, CSS, data communication based on http protocol of RESTful style, and used the JSON data format for data transfer. The server uses programming language Golang that Google developed with the characteristics of compiled, parallelized and garbage collection. Web API layer is based on RESTful Web API open source framework Martini that can be divided to three layers of routing management, controller and RPC proxy. The route management module is responsible for routing resource allocation; the controller module defines the behavior of the application system, and exposure resources to the front page provided by the HTTP method; RPC proxy module is responsible for communication with the model layer, using MongoDB stored data and highcharts.js show the fault data.

Implementation results are shown in Figure 5, Figure 6 and Figure7.
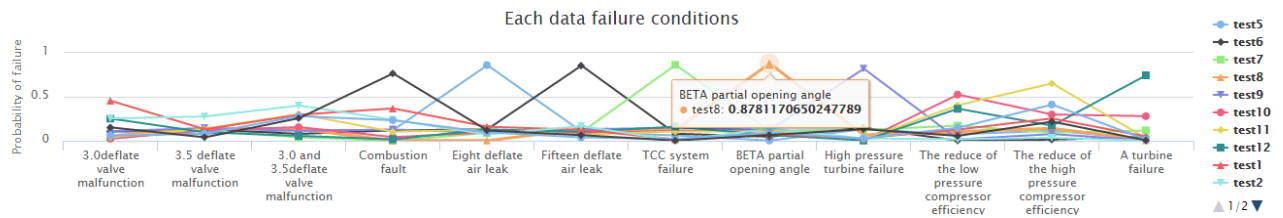
Neural Network Fault Analysis:

Figure 5. Visualizations of engine fault status diagnostics data -Neural Network
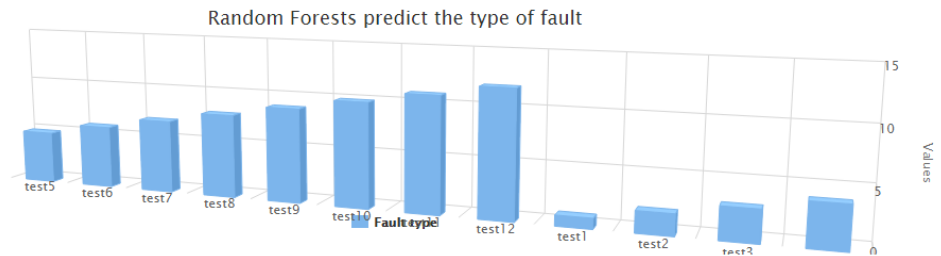
Random Forests Fault Analysis:

Figure 6. Visualizations of engine fault status diagnostics data -Random Forest
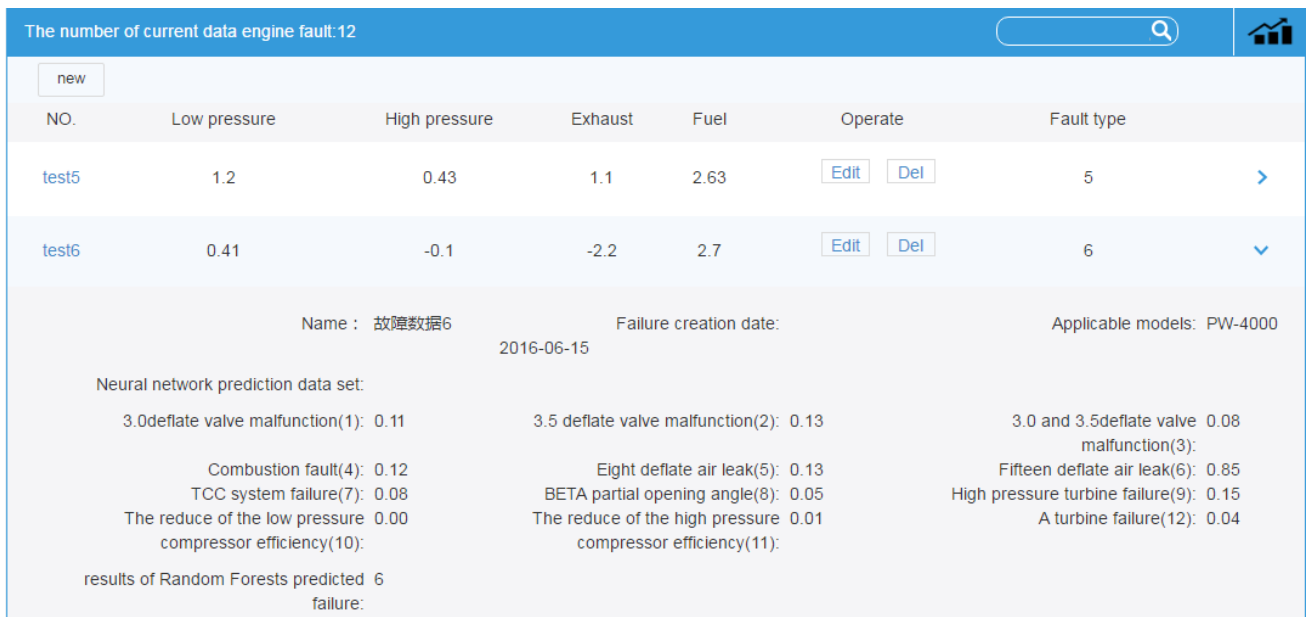

Figure 7. Diagnosis results of engine fault status

5 Conclusion

The paper deals with the research and application of BP Neural Network algorithm and Random Forest algorithm in aircraft engine fault diagnosis, the effectiveness of the method is verified by experiments, and realize the corresponding web application.The experimental results show that both methods have their advantages and disadvantages. After combining the two aircraft engine fault diagnosis algorithm,we can achieve the desired results that more accurate type of engine fault and potential problems can be discovered.

References

[1] Li Ai.Research on Key Techniques of Aeroengine Wear Fault Intelligent Diagnosis [D]. Nanjing University of Aeronautics and Astronautics, 2013.

[2] Tao DongQi, Bao CuiMei, Yi-hui.Fault diagnosis for transformer based on Random Forest [J] electronic devices, 2015,04: 840-844.

[3] Wu YueWei. Study Of Aircraft Engine Fault Diagnosis Based On RBF Neural Network [D]. China Civil Aviation College, 2006.

[4] Review of machine learning algorithms -Random Forest [EB / OL] http://www.cnblogs.com/wentingtu/archive/2011/12/22/2297405.html

[5] Yan RenWu, Ye QingZhou, Zhou Li.Application of Random Forest algorithm to fault diagnosis of power electronic circuit. [J] Wuhan University (Engineering Science), 2013,06: 742-746.