Research on the Analysis Method of Engine Assembly Quality Based on Data Driven

DOU Xin¹, XU Jia-li², WU Jia-jing³, JIA Dong-ning*¹,²

¹School of Information Science and Engineering, Ocean University of China, Qingdao, 266100, China;
²Pilot National Laboratory for Marine Science and Technology, Qingdao, 266237, China.

*Correspondence: jiadn@ouc.edu.cn

Abstract. In the context of the country’s vigorous development of the industrial Internet, the role of big data has become increasingly prominent. Through the mining and analysis of the value of industrial big data, the efficiency and quality of industrial production will be effectively improved. However, the current research lacks direct quantitative research on the correlation between bolt tightening position data and engine quality. Based on the data of a certain production line of Weichai, the SimpleKMeans clustering and Apriori correlation analysis model for the assembly process data and product quality of a certain type of diesel engine is constructed based on the Weka tool. The mining results of the association rule between the bolt tightening data of the eight stations and the cylinder head vibration peak data show that the cylinder head main bolts, cylinder head auxiliary bolts, camshaft gear bolts and connecting rod bolts are closely related to the cylinder head vibration data, and the confidence levels are respectively 0.78, 0.70, 0.63, 0.52. Therefore, when the engine vibration is abnormal, the tightening data of the cylinder head main bolts, cylinder head auxiliary bolts, camshaft gear bolts and connecting rod bolts are detected first to effectively improve the quality and efficiency of the assembly process.

1. Introduction

Since the State Council issued the "Guiding Opinions on Deepening the "Internet + Advanced Manufacturing" Development of the Industrial Internet" in November 2017, the construction of the domestic industrial Internet has developed rapidly [1]. Simultaneously, the Industrial Internet, one of the seven major areas of the country’s new infrastructure, has been supported and strengthened by policies issued in many places, achieving a high degree of integration of industrialization and informatization. Machine manufacturing is an important part of the manufacturing industry and one of the key targets of the new infrastructure industrial Internet. As a power engine, the engine is widely used in construction machinery, vehicles, generator sets and other fields, and its performance and quality have always been research hot spots [2][9]. Meanwhile, diesel engines have a wide range of applications, considerable market demand, and higher requirements for fuel consumption, temperature and pressure resistance, component stiffness and accuracy [10][13], so diesel engines have greater research value.

Current research [14][18] mainly uses assembly data to explore the law of action between different parameters in the assembly process, and lacks a direct quantitative study of the relationship between tightening data and engine quality. Additionally, the research involving bolt tightening data does not consider the factor of bolt position. Therefore, based on the Weka tool, this paper uses SimpleKMeans and Apriori two algorithms to comprehensively analyze the clustering and correlation between bolt
tightening data and cylinder head vibration data of different stations. By constructing the bolt tightening torque and tightening angle of the eight stations of camshaft gear bolts, intermediate gear fastening bolts, flywheel bolts, flywheel housing bolts, connecting rod bolts, cylinder head main bolts, cylinder head auxiliary bolts, and main bearing bolts, multiple data mining models between the data and the cylinder head vibration peak data realize the accuracy of the bolt tightening data range and the quantitative association mining of the relationship between the bolt tightening data and the cylinder head vibration data, thereby improving the quality and efficiency of engine assembly.

2. Methods and tools

2.1. K-means

The K-means algorithm is a clustering algorithm that divides the sample set according to the number k of clusters specified by the user. The division strategy is based on minimizing the loss function, which is solving the optimization problem of Equation (1):

$$C^* = \arg \min_C \sum_{i=1}^{k} \sum_{C(i)=l} ||x_i - \bar{x}_l||^2$$  \hspace{1cm} (1)

among them C(i), l ∈ {1, 2, ⋯, k} is the partition function of the ith cluster, \(\bar{x}_l = (\bar{x}_{l1}, \bar{x}_{l2}, \cdots, \bar{x}_{lm})^T\), l ∈ {1, 2, ⋯, k} is the mean of theith cluster, and \(x_i\) is the feature vector in the sample set.

The K-means algorithm is actually a two-step iterative process. First, select k cluster centers, assign each sample to the nearest cluster center, and assign samples to the same cluster center to form a cluster; then, update the cluster center according to the mean value of each cluster sample. Repeat the above two processes until the clusters and cluster centers no longer change.

The calculation of similarity in the K-means algorithm usually uses Euclidean distance. Euclidean distance refers to the straight-line distance between two points in Euclidean space. The calculation formula between two points in two-dimensional space is Equation (2). In addition, Sum of Squared Error (SSE) is usually used as a standard to measure the quality of clustering, and its calculation formula is Equation (3), among them \(c_i\) is the cluster center of the ith cluster.

$$\text{dist}(x, y) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$  \hspace{1cm} (2)

$$\text{SSE} = \sum_{l=1}^{k} \sum_{x \in c_l} \text{dist}(c_l, x)^2$$  \hspace{1cm} (3)

The SimpleKMeans algorithm provided by Weka’s cluster interface uses the K-means clustering algorithm, in which the number of clusters is given by the setting of the parameter num clusters; the setting of the parameter seed can be the k cluster centers of the first operation for reference. The performance evaluation standard of SimpleKMeans algorithm adopts the SSE. The smaller the value, the better the clustering result.

2.2. Apriori

Apriori algorithm is the first association rule mining algorithm. Its algorithm mainly uses the idea of layer-by-layer iteration to complete the two steps of generating frequent item sets and mining strong association rules. Generating frequent item sets includes two important operations, which are connection and pruning. The specific execution steps of the algorithm are as follows:

The first step is to find all frequent item sets, which meet the support not less than the predetermined minimum support threshold:

1. Scan the data set and calculate the degree of support. The Support of the association rule can reflect the importance of the association rule in the transaction set, and the calculation formula is as Equation (4), where \(A \rightarrow B(A \subseteq I, B \subseteq I, A \cap B = \emptyset)\), and \(I = \{i_1, i_2, \cdots, i_m\}\) is item set.
(2) Generate frequent item sets. By retaining the item sets that are not less than the minimum support threshold in the candidate item sets, new frequent item sets $L_k$ can be generated.

(3) Connection. If $L_{k-1}(k>2)$ meet the connection judgment condition, the frequent item set $L_{k-1}$ connect with itself to generate a new candidate set $C_k$. $L_{k-1}$ need to meet Equation (5) to perform the connection operation, among them $I_1$ and $I_2$ are two subsets of the $(k-1)$ item sets, and $I_j[j]$ represents jth in the item sets $I_j$.

$$((I_1[1]=I_2[1]) \land (I_1[2]=I_2[2]) \land \ldots \land (I_1[k-2]=I_2[k-2]) \land (I_1[k-1] \neq I_2[k-1])) \neq \emptyset$$

(4) Pruning. Using the nature of the Apriori algorithm which is "all subsets of frequent item sets are frequent", candidate set $C_i$ which is not in the frequent item sets $L_{k-1}$ delete directly to avoid subsequent unnecessary support calculation operations.

(5) Generate a set of candidates. The candidate item sets $C_i$ refers to the item sets obtained from the top-down merge operation in the connection step of the Apriori algorithm.

Repeat the operations from (1) to (5) until no new candidate item sets can be generated, and the frequent item sets generation algorithm ends.

The second step is to mine strong association rules, which are association rules whose support is greater than the preset minimum support threshold and whose confidence is not less than the specified minimum confidence threshold. Among them, the Confidence refers to the probability that B is derived from $A \rightarrow B$ under the premise of the occurrence of the prerequisite A, which reflects the credibility of the association rule, and is calculated as Equation (6).

$$\text{Confidence}(A \rightarrow B) = \frac{P(A \cap B)}{P(A)}$$

Weka's Associate interface provides Apriori algorithm for association rule mining. The minimum support and minimum confidence thresholds can be set through the lowerBoundMinSupport and minMetric parameters. The performance metrics in the Apriori algorithm are based on confidence, lift, leverage, and certainty [22]. The Lift calculation method is as Equation (7). If Lift>1, the A and B sets have a strong correlation, and the larger the value, the stronger the correlation; the Leverage calculation method is as Equation (8). The smaller the value, the weaker the relationship; the calculation method of Conviction is as Equation (9), where $\overline{B}$ means that B does not appear or occurs, and its value is proportional to the close relationship.

$$\text{Lift}(A \rightarrow B) = \frac{P(B | A)}{P(B)}$$

$$\text{Leverage}(A \rightarrow B) = P(A \cap B) - P(A).P(B)$$

$$\text{Conviction}(A \rightarrow B) = \frac{P(A).P(\overline{B})}{P(A \cap \overline{B})}$$

$2.3.\text{Weka} \text{ tool}$

Weka [23] (Waikato Environment for Knowledge Analysis) is an open source machine learning and data mining tool developed by Waikato University in New Zealand. It has a development history of more than 20 years and is currently one of the most complete data mining tools. Firstly, Weka supports JAVA language for knowledge discovery and secondary development; secondly, the tool integrates basic machine learning algorithms, including classification, clustering, association analysis, and so on, and
provides preprocessing and visualization interfaces to assist users in operations; additionally, the software provides multiple functional operation portals such as explorer interface, knowledge flow interface, experimenter interface and command line interface for users to choose. Weka has complete functions and simple operation. Users only need to select the algorithm and set the parameters to easily achieve the effect.

Weka supports file input in ARFF (Attribute Relation File Format) and CSV format, and also supports access to the database through JDBC. Among them, ARFF is an ASCII text file presented in the form of a two-dimensional table. And provide support for five attribute formats: string, relational, nominal, numeric, and date.

The version of the Weka tool used in this article is Weka-3.8.4, which was updated in 2016. This version has complete functions and stable performance.

3. Experiment

3.1. Experimental data
The data used in this paper are 800 pieces of bolt tightening data and cylinder head vibration data from a certain model of diesel engine assembly process. Specifically, the eight-position bolts of the camshaft gear bolts, intermediate gear fastening bolts, flywheel bolts, flywheel housing bolts, connecting rod bolts, cylinder head main bolts, cylinder head auxiliary bolts, and main bearing bolts during the last tightening process. Tightening torque data, tightening angle data and corresponding cylinder head vibration peak detection data.

In addition, Table 1 gives the tightening standard data corresponding to the bolt tightening station of a certain type of diesel engine studied in this paper.

| Bolt name               | Thread specification | Tightening torque /N.m | Tightening angle /° |
|-------------------------|----------------------|------------------------|---------------------|
| Camshaft gear bolt      | M8                   | 31-33                  | 0                   |
| Intermediate gear       | M10                  | 55-65                  | 85-95               |
| Fastening bolt          | M12                  | 40-60                  | 115-125             |
| Flywheel bolt           | M14×1.5              | 55-80                  | 85-95               |
| Flywheel housing bolt   | M12                  | 60-80                  | 85-95               |
| Connecting rod bolt     | M14×1.5              | 119-121                | 85-95               |
| Cylinder head main     | M16                  | 200-210                | 85-95               |
| Bolt                    |                      |                        |                     |
| Cylinder head auxiliary | M12                  | 100-110                | 85-95               |
| Main bearing bolt       | M18                  | 250-275                | 0                   |

3.2. Pretreatment
First, organize the data into two CSV files, one of which includes bolt tightening torque data and corresponding cylinder head vibration peak data for eight stations, and the other includes tightening angle data and cylinder head vibration data; The file is preprocessed. Because the tightening angle data of the camshaft gear bolt and the main bearing shaft bolt are both 0, the corresponding data is deleted; finally, the two CSV files are converted into ARFF files through the Weka tool.

3.3. Data mining
First, load the data into the Weka tool, and perform simple processing and analysis through the preprocessing interface; then, enter the clustering window, select the SimpleKMeans algorithm and set the random seed to 40, set the number of clusters to 8, and get the value Clustering results of type data; finally, return to the preprocessing interface to discretize the data using NumericToNominal filter, and enter the correlation interface, through Apriori -N 10 -T 0 -C 0.3 -D 0.05 -U 1.0 -M 0.001 -S 1.0 -e -1 is set for association rule mining. Perform the above operations on the tightening torque data file and the tightening angle data file in turn.
4. Results

4.1. Analysis of preprocessing results
The distribution diagrams of bolt tightening torque and tightening angle data are shown in Figure 1 and Figure 2, respectively:

(a) Camshaft gear bolts; (b) Intermediate gear fastening bolts; (c) Flywheel bolts; (d) Flywheel housing bolts; (e) Connecting rod bolts; (f) Cylinder head main bolts; (g) Cylinder head Auxiliary bolt; (h) main bearing bolt

Figure 1. Distribution graph of tightening torque data.

(a) Intermediate gear fastening bolts; (b) Flywheel bolts; (c) Flywheel housing bolts; (d) Connecting rod bolts; (e) Cylinder head main bolts; (f) Cylinder head auxiliary bolts

Figure 2. Distribution graph of tightening angle data.
From the results of preprocessing experiments, it is found that the data are distributed in the standard data reference range, but the degree of distribution and concentration varies. According to the distribution characteristics of bolt tightening data, the tightening reference data in Table 1 can be refined into Table 2.

Table 2. Reference table of precise data for bolt tightening.

| Bolt name              | Thread specification | Tightening torque /N.m | Tightening angle /° |
|------------------------|----------------------|------------------------|---------------------|
| Camshaft gear bolt     | M8                   | 31.8-32.3              | 0                   |
| Intermediate gear fastening bolt | M10             | 55-60                  | 89-92               |
| Flywheel bolt          | M14×1.5              | 60-70                  | 90-95               |
| Flywheel housing bolt  | M12                  | 40-50                  | 120-125             |
| Connecting rod bolt    | M14×1.5              | 119.7-120.4            | 88-93               |
| Cylinder head main bolt| M16                  | 205-210                | 85-88               |
| Cylinder head auxiliary bolt | M12         | 105-110                | 85-88               |
| Main bearing bolt      | M18                  | 260-265                | 0                   |

Through preprocessing analysis, on the one hand, the reference range of the bolt tightening data in Table 1 can be reduced. The tightening torque data can be reduced to 20% of the range threshold and the tightening angle is 30%; on the other hand, flywheel bolts and flywheels can also be found. The tightening data distribution characteristics of the shell bolts, cylinder head main bolts and cylinder head auxiliary bolts are similar. In the bolt tightening torque data, 88.1% of flywheel bolts and flywheel housing bolts are mainly distributed in the first half of the range of Table 1, and 87.5% of the data of cylinder head main bolts and cylinder head auxiliary bolts are concentrated in the second half of the range of Table 1. In bolts In the tightening angle data, 87.3% of the data of flywheel bolts and flywheel housing bolts are gathered in the second half of the range of Table 1, and 89% of the data of cylinder head main bolts and cylinder head auxiliary bolts are concentrated in the first half of the range of Table 1.

4.2. Cluster analysis

Table 3 and Table 4 are the clustering results of bolt tightening torque data and tightening angle data, respectively. The analysis of the two sets of results is as follows:

First, the core clusters in the two sets of results can reflect the overall characteristics of the clustering results. In the tightening torque results, cluster 2, cluster 3, cluster 6, and cluster 7 account for 73% of the total. The vibration data of these four clusters are all within 2.59 μm, which is only 0.02 μm away from the average value; The ratio of the difference between the data of the eight stations in the cluster and the mean value of the sample to the reference range is within 23%, and the bolt tightening torque data of the three stations of the cylinder head main bolt, auxiliary bolt and camshaft gear are approximately the same as the sample average. In the tightening angle results, cluster 3 and cluster 5 account for 62%. The tightening angle data of the six stations in the cluster are all within the precise range of Table 2, and the corresponding vibration data is less than the sample average, reflecting the engine's The angle tightening data is of high quality.

Second, there are clusters with large deviations from the sample mean in both sets of results. In the tightening torque results, the difference between the tightening torques of the cylinder head main bolts and auxiliary bolts of clusters 0 and 4 and the sample average is as high as 4 Nm. At the same time, these two parameters deviate from the precise range of Table 2, which eventually leads to The vibration peaks of these two clusters are 0.9 μm higher than the sample average. In the tightening angle results, the tightening angle data of the cylinder head main and auxiliary bolts of clusters 1 and 4 have a deviation of 5.54° from the sample average. In addition, the data of these two stations are beyond the precise range provided in Table 2, based on the above two This is why it deviates from the sample cluster center.

Finally, the remaining clusters are between the first two cases. In the tightening torque results, cluster 1 and cluster 5 did not simultaneously show that the main and auxiliary bolts of the cylinder head
deviated from the precise range. The number of deviation stations was only one on average; in the tightening angle results, cluster 0, cluster 2, cluster 6, and cluster 7 in the four clusters, the data of two stations on average deviated from the accurate data reference range, and the bolt tightening angle data of the main and auxiliary stations of the cylinder head did not appear in the four clusters at the same time. The final result is the cylinder head vibration data. The deviations from the sample mean are all within 0.33μm.

Table 3. Result table of tightening torque clustering.

| Cluster | Camshaft gear bolt /N.m | Intermediate gear fastening bolt /N.m | Flywheel bolt /N.m | Flywheel housing bolt /N.m | Connecting rod bolt /N.m | Cylinder head main bolt /N.m | Cylinder head auxiliary bolt /N.m | Main bearing bolt /N.m | Cylinder head vibration peak /μm | Percent -age /% |
|---------|--------------------------|--------------------------------------|-------------------|---------------------------|--------------------------|----------------------------|-------------------------------|------------------------|-----------------------------|---------------------|
| Cluster 0 | 32.0159 | 59.4135 | 68.527 | 52.673 | 120.0016 | 202.0243 | 103.0243 | 263.4108 | 3.4192 | 5 |
| Cluster 1 | 31.9823 | 58.0021 | 68.7234 | 50.5532 | 120.0326 | 207.1021 | 103.9532 | 256.766 | 2.7326 | 6 |
| Cluster 2 | 32.044 | 57.127 | 62.3794 | 46.6246 | 120.0689 | 206.7974 | 106.2548 | 261.0714 | 2.43 | 16 |
| Cluster 3 | 32.067 | 57.473 | 66.3284 | 45.7532 | 120.0285 | 206.917 | 108.1404 | 253.756 | 2.2952 | 18 |
| Cluster 4 | 32.067 | 57.473 | 66.3284 | 45.7532 | 120.0285 | 206.917 | 108.1404 | 253.756 | 2.2952 | 18 |
| Cluster 5 | 32.044 | 56.6512 | 70.0643 | 50.1536 | 120.0094 | 207.0796 | 107.1747 | 257.7095 | 2.4713 | 11 |
| Cluster 6 | 32.047 | 57.8857 | 68.4048 | 47.7131 | 120.0377 | 206.5488 | 108.1571 | 264.9089 | 2.5853 | 21 |
| Cluster 7 | 32.0412 | 57.6643 | 65.672 | 42.258 | 120.0999 | 206.5594 | 108.2664 | 262.0538 | 2.4748 | 18 |
| Sample mean | 32.0416 | 58.4826 | 66.8169 | 46.8198 | 120.0558 | 206.3745 | 107.3212 | 260.7562 | 2.5706 | 100 |

Table 4. Result table of tightening angle clustering.

| Cluster | Intermediate gear fastening bolt /° | Flywheel bolt /° | Flywheel housing bolt /° | Connecting rod bolt /° | Cylinder head main bolt /° | Cylinder head auxiliary bolt /° | Cylinder head vibration peak /μm | Percent -age /% |
|---------|------------------------------------|-----------------|-------------------------|-------------------------|---------------------------|-------------------------------|-----------------------------|---------------------|
| Cluster 0 | 87.4524 | 87.981 | 118.8595 | 89.1762 | 87.5833 | 88.5452 | 2.8321 | 5 |
| Cluster 1 | 87.0811 | 91.2108 | 122.1108 | 89.6946 | 92.8514 | 92.2081 | 3.1559 | 5 |
| Cluster 2 | 90.6983 | 92.1492 | 120.8787 | 90.6542 | 87.4373 | 93.4373 | 2.9012 | 7 |
| Cluster 3 | 90.5215 | 90.812 | 122.3402 | 89.9386 | 86.561 | 86.602 | 2.3781 | 31 |
| Cluster 4 | 91.28 | 90.52 | 120.5133 | 90.3578 | 92.9756 | 91.1933 | 3.0302 | 6 |
| Cluster 5 | 90.3122 | 93.6065 | 122.0722 | 90.0596 | 86.6637 | 86.6363 | 2.4502 | 31 |
| Cluster 6 | 90.9481 | 91.8304 | 122.5722 | 89.9494 | 87.0924 | 90.1975 | 2.5682 | 10 |
| Cluster 7 | 89.9238 | 89.3548 | 116.8952 | 89.731 | 86.8095 | 88.05 | 2.4682 | 5 |
| Sample mean | 90.2036 | 91.6439 | 121.5913 | 89.9909 | 87.4343 | 88.1519 | 2.5706 | 100 |

Through the analysis of the tightening torque and tightening angle cluster results, it can be found that at least 62% of the data are distributed within the precise range of Table 2, which provides evidence for the reliability of the preprocessing results; at the same time, through the comparison between different clusters, it can be Predict the relationship between different stations. For example, through clusters 0 and 4 in Table 3, and clusters 1 and 4 in Table 4, it can be predicted that the main and auxiliary bolts of the cylinder head may be related to each other, and the two have a relatively large impact on the vibration of the cylinder head.

4.3 Association result analysis

The mining results of 10 association rules between bolt tightening torque data and cylinder head vibration data are shown in Table 6. Analysis of the results can be obtained: On the one hand, the bolt tightening torque data of the four stations of the cylinder head main bolt, cylinder head auxiliary bolt, camshaft gear bolt and connecting rod bolt have a strong correlation with the peak vibration data of the cylinder head. Among them, the correlation between cylinder head vibration data and the tightening torque of the cylinder head main bolt is the strongest, with a confidence of 0.79; the correlation between
connecting rod bolts and cylinder head vibration is weak, with a confidence of only 0.43. On the other hand, there are also certain correlations between flywheel bolts and flywheel housing bolts, and between cylinder head main bolts and cylinder head auxiliary bolts. The confidence levels of the association rules are 0.56 and 0.38, respectively.

The results of the association rule mining between the bolt tightening angle data of the six stations and the cylinder head vibration data are shown in Table 7. The correlation results can be found: First, the cylinder head main bolts, cylinder head auxiliary bolts and connecting rod bolts are tightened. The angle data still has a strong correlation with the cylinder head vibration data, and the confidence level is between 0.6 and 0.77; in addition, the bolt tightening between the flywheel bolt and the flywheel housing bolt, and the cylinder head main bolt and the cylinder head auxiliary bolt still exist. For the association relationship, the confidence levels of the association rules are 0.5 and 0.48 respectively. The results of the association rule mining of bolt tightening angle data verify the reliability of the tightening torque data results.

According to the mining results of the above association rules, the bolt positions that affect the vibration of the cylinder head can be sorted. The sorting results are shown in the Table 8.

Therefore, when the engine cylinder head vibration data is abnormal, the tightening data of the cylinder head main bolts and cylinder head auxiliary bolts should be checked first. If there is no quality problem, then check the tightening quality of the camshaft gear bolts and connecting rod bolts, and so on. Find the cause of the abnormality in a limited time and save the time required for the abnormal data investigation process. In addition, there is also a certain correlation between the flywheel bolts and flywheel housing bolts, and between the cylinder head main bolts and cylinder head auxiliary bolts. Therefore, when the data of one of the stations is abnormal, the data of the other station can be verified first. To maximize the assembly quality and efficiency of the engine.

Table 5. Description table of name and code.

| Name | Code |
|------|------|
| Camshaft gear bolt | A |
| Intermediate gear fastening bolt | B |
| Flywheel bolt | C |
| Flywheel housing bolt | D |
| Connecting rod bolt | E |
| Cylinder head main bolt | F |
| Cylinder head auxiliary bolt | G |
| Main bearing bolt | H |
| Cylinder head vibration peak | I |
| Bolt tightening torque | 1 |
| Bolt tightening angle | 2 |

Table 6. Result table of tightening torque association rule mining.

| ID | Association rules | Confidence | Lift | Leverage | Certainty |
|----|------------------|------------|------|----------|-----------|
| 1  | I=2.68==>F I=205.4 | 19/24=0.79  | 19.79 | 0.02     | 3.84      |
| 2  | I=2.44==>G I=108.8 | 22/33=0.67  | 11.85 | 0.03     | 2.6       |
| 3  | I=2.59==>A I=32.29 | 10/16=0.63  | 21.74 | 0.01     | 2.22      |
| 4  | A=205.4===>I=2.68  | 19/32=0.59  | 19.79 | 0.02     | 2.22      |
| 5  | C=65.5===>D I=48.7 | 9/16=0.56   | 18.75 | 0.01     | 1.94      |
| 6  | G=108.8===>I=2.44  | 22/45=0.49  | 11.85 | 0.03     | 1.8       |
| 7  | A=32.29===>I=2.59  | 10/23=0.43  | 21.74 | 0.01     | 1.61      |
| 8  | E=120.39===>I=2.37  | 9/21=0.43   | 12.24 | 0.01     | 1.56      |
| 9  | D=48.7===>C I=65.5 | 9/24=0.38   | 18.75 | 0.01     | 1.47      |
| 10 | G=107.9===>F I=205.8| 9/24=0.38   | 12.5  | 0.01     | 1.46      |

Table 7. Result table of tightening angle association rule mining.

| ID | Association rules | Confidence | Lift | Leverage | Certainty |
|----|------------------|------------|------|----------|-----------|
| 1  | I=2.68==>F 2=86.7| 23/30=0.77  | 16.58 | 0.03     | 3.58      |
| 2  | I=2.44==>G 2=86.6| 19/26=0.73  | 17.72 | 0.02     | 3.12      |
| 3  | F 2=86.7===>I=2.68| 23/37=0.62  | 16.58 | 0.03     | 2.37      |
| 4  | I=2.37==>E 2=90.8| 9/15=0.6    | 22.86 | 0.01     | 2.09      |
| 5  | G 2=86.6===>I=2.44| 19/33=0.58  | 17.72 | 0.02     | 2.13      |
| 6  | G 2=87.5===>F 2=87.9| 9/18=0.5    | 14.29 | 0.01     | 1.74      |
Table 8. Credibility result table of association rules included bolt station and cylinder head vibration.

| Bolt name          | Connecting rod bolt | Camshaft gear bolt | Cylinder head main bolt | Cylinder head auxiliary bolt |
|--------------------|---------------------|--------------------|-------------------------|-----------------------------|
| Confidence         | 0.515               | 0.63               | 0.7                     | 0.78                        |

5. Conclusion

Aiming at the lack of influence of bolt tightening station factors on engine quality data in the engine assembly research, this paper proposes a systematic mining method for bolt tightening and cylinder head vibration data at eight stations. This method is based on Weka tools and constructs SimpleKMeans and Apriori algorithm models, which realizes the reduction of the tightening reference data range, the correlation analysis of tightening and quality data, and ensures the quality and efficiency of engine assembly. The specific experimental results are described as follows:

1. Through the visual analysis of the preprocessed data, the tightening reference range in Table 1 is reduced to the tightening accuracy range in Table 2. The tightening torque data range can be reduced by up to 70%, and the tightening torque data can be reduced by up to 80%.

2. On the one hand, the clustering results can verify the validity of the preprocessing analysis results, on the other hand, it can also discover the concentrated characteristics of engine assembly data and the proportion of outlier data. Among them, 73% of the tightening torque data and 62% of the tightening angle data have high quality requirements, and 5% of the tightening torque and tightening angle data are on the margin of qualification.

3. Association rule mining experiment results show that cylinder head vibration is closely related to the four stations of cylinder head main bolts, cylinder head auxiliary bolts, camshaft gear bolts and connecting rod bolts, with a confidence level of up to 0.78; in addition, there is also a certain correlation between the cylinder head main bolt and cylinder head auxiliary bolt, and between the flywheel bolt and the flywheel housing bolt, and the confidence levels are 0.5 and 0.69, respectively.

Acknowledgments

This research was funded by National Key R&D Program of China named manufacturing big data driven whole process production operation optimization and decision technology project, which grant number is 2018YFB1703104.

References

[1] SUN Bo-yu. Research on Industrial Policy Issues of "Internet +" Advanced Manufacturing[D]. Dalian: School of Public Administration, Dongbei University of Finance and Economics, 2018: 16-29.

[2] Bhowmik S, Paul A, Panua R, et al. Artificial intelligence based gene expression programming (GEP) model prediction of Diesel engine performances and exhaust emissions under Diesosenol fuel strategies[J]. Fuel, 2019, 235: 317-325.

[3] Yang J, Dong X, Wu Q, et al. Effects of enhanced tumble ratios on the in-cylinder performance of a gasline direct injection optical engine[J]. Applied energy, 2019, 236: 137-146.

[4] Ağbulut Ü, Karagöz M, Sarıdemir S, et al. Impact of various metal-oxide based nanoparticles and biodiesel blends on the combustion, performance, emission, vibration and noise characteristics of a CI engine[J]. Fuel, 2020, 270: 117-221.

[5] Da Lio M, Bortoluzzi D, Rosati Papini G P. Modelling longitudinal vehicle dynamics with neural networks[J]. Vehicle System Dynamics, 2019: 1-19.

[6] Lan S, Smith A, Stobart R, et al. Feasibility study on a vehicular thermoelectric generator for both
waste heat recovery and engine oil warm-up[J]. Applied Energy, 2019, 242: 273-284.

[7] Fonseca J H, Han G, Quagliato L, et al. Design and numerical evaluation of recycled-carbon-fiber-reinforced polymer/metal hybrid engine cradle concepts[J]. International Journal of Mechanical Sciences, 2019, 163: 105-115.

[8] Chen W, Pan J, Liu Y, et al. Numerical investigation of direct injection stratified charge combustion in a natural gas-diesel rotary engine[J]. Applied Energy, 2019, 233: 453-467.

[9] Patidar S K, Raheman H. Performance and durability analysis of a single-cylinder direct injection diesel engine operated with water emulsified biodiesel-diesel fuel blend[J]. Fuel, 2020, 273: 117-179.

[10] Wang S, Zhu R. Study on load sharing behavior of coupling gear-rotor-bearing system of GTF aero-engine based on multi-support of rotors[J]. Mechanism and Machine Theory, 2020, 147: 103-164.

[11] Zhao B, Xie L, Song J, et al. Fatigue life prediction of aero-engine compressor disk based on a new stress field intensity approach[J]. International Journal of Mechanical Sciences, 2020, 165: 105-190.

[12] Ganesh D, Ayyappan P R, Murugan R. Experimental investigation of iso-butanol/diesel reactivity controlled compression ignition combustion in a non-road diesel engine[J]. Applied Energy, 2019, 242: 1307-1319.

[13] Guerrero D P, Jiménez-Espadafor F J. Torsional system dynamics of low speed diesel engines based on instantaneous torque: Application to engine diagnosis[J]. Mechanical Systems and Signal Processing, 2019, 116: 858-878.

[14] JIANG Yun-fan, LIAO Ming-fu, CHEN Jing, et al. Support Stiffness Dependent on Aero-engine Bearing Assembly Conditions[J]. Journal of Vibration, Measurement and Diagnosis, 2020,40(02): 348-354+421

[15] LIU Zhong-hua, JIA Duo,WANG Xin, et al. Clamp assembly stress test of aero-engine and assembly parameter control method[J]. Journal of Aerospace Power, 2020,35(02):368-377.

[16] KANG Na, WANG Ping-an, DAI Wei, et al. Problem solving process of spark plug hole sealing leakage during engine assembly[J]. Manufacturing Technology & Machine Tool, 2019(11):152-155.

[17] CHEN Zhi-ying, LIU Honglei, ZHOU Ping. Robustness Optimization of Dynamic Assembly Parameters for Aero-Engine Spline[J]. Journal of Propulsion Technology, 2018,39(01):160-168.

[18] SHANGGUAN Wan-bin, DU Xiao-ze, NIE Jun, et al. Optimization of Drum-Shape Design and Assembly Performance of Rubber Torsional Vibration Absorber for Engine Crankshafts[J]. Chinese Internal Combustion Engine Engineering, 2017,38(03):149-156.

[19] Hartigan J A, Wong M A. Algorithm AS 136: A K-means clustering algorithm[J]. Journal of the royal statistical society. series c (applied statistics), 1979, 28(1): 100-108.

[20] Sudarsan V, Sugumar R. Building a distributed K-Means model for Weka using remote method invocation (RMI) feature of Java[J]. Concurrency and Computation: Practice and Experience, 2019, 31(14): e5313.

[21] Agarwal R, Srikan R. Fast algorithms for mining association rules[C]//Proc. 20th int. conf. very large data bases,VLDB. Chile,1994: 487-499.

[22] Gao F, Khandelwal A, Liu J. Mining Frequent item sets Using Improved Apriori on Spark[C] //Proceedings of the 2019 3rd International Conference on Information System and Data Mining. Houston,2019: 87-91.

[23] Markov Z, Russell I. An introduction to the WEKA data mining system[J]. ACM SIGCSE Bulletin, 2006, 38(3): 367-368.