Research on Assembly Line Optimization Based on Machine Learning

Zhang Peng
School of Mechatronic Engineering
Xi'an Technological University
Xi'an, China
E-mail: 1101693806@qq.com

Fang Yadong
School of Mechatronic Engineering
Xi'an Technological University
Xi'an, China
E-mail: tomfangok@126.com

Liang Xiaowei
School of Mechatronic Engineering
Xi'an Technological University
Xi'an, China
E-mail: doris_liang_@outlook.com

Chen Jingwen
School of Mechatronic Engineering
Xi'an Technological University
Xi'an, China
E-mail: chenjingwen6539@163.com

Abstract—**How to conduct a wide range of problem analysis on production line work in a short time and standardize, its work has become a major difficulty for enterprises to improve work efficiency. This paper takes M company's Rail car assembly as an example to conduct the job optimization. It uses the k-means algorithm in machine learning to conduct clustering analysis on job data，Identify common and unusual factors in the assignment. Establishing different text dictionaries aims to normalize the expression of texts and help identify as well as remove non - value-added work efficiently. Application of machine learning makes it easy for management personnel to identify the operational bottlenecks of the entire production line, achieve standardized operations, and improve production efficiency.**

Keywords-K-means Algorithm; Text processing; job standardization; Non-value-added Operations

I   INTRODUCTION

With the advent of the era of big data, machine learning technology has developed rapidly and is gradually applied to various fields. Informatization means to stimulate and promote the leap-forward development of industrialization, while bringing considerable benefits to enterprises. The integration of the two industries brings challenges to enterprises while also bringing challenges to many enterprises. Traditional high-handed manufacturing industry is getting out of trouble, proposing to improve the company's current business and production management site by means of information technology, implementing lean production and standardize operations.

As one of the machine learning methods, K-means Algorithm can realize data mining. The purpose of clustering is to cluster samples based on the similarity of data samples, to minimize the similarity between similar samples, and to increase the similarity between different samples. This article takes M company rail car manufacturing company as an example. According to the current situation of the assembly workshop and the characteristics of the operation, K-means Algorithm is used to cluster and analyze the realistic data of job. Finding out the whole production line value-added operations in a short time, then enterprises can take improvement methods more conveniently and meet the requirements of standardized operations.

II   SITUATION ANALYSIS

M Company is a rail car manufacturing company, and the assembly workshop belongs to the high manual operation workshop. In the first assembly workshop, a station-based production line was established. During the trial operation, there were various problems including material distribution, quality inspection, on-site visualization, process standard operation and production line balance rate. The problems were randomly found in each operation process. The production line consists of a total of 6 stations, a total of 180 large and small operations and 5,400 operation steps. Analyzing and improving each process is not only a huge workload, but also cannot effectively judge the common problem factors in each process operation. Result in repeated improvement, invalid improvement and other results.

With the huge amount of process operation and the random existence of various production problems in each process operation. In order to help enterprises solve the dilemma in the shortest time, we use the machine learning K-Means Algorithm to conduct clustering analysis on realistic data of production line operations, dig out the common factors in the production line operation and then conduct unified processing on these common factors. The specific analysis flow chart is shown in Figure 1.

Figure 1. Analysis flow char

III PRODUCTION LION OPERATION REALISM

Select 300 operation steps in 10 processes of the production line as the analysis object, using stopwatch timing and site recording for operation realism in each process. The realistic operation table (headlamp installation) is as shown in Table 1.
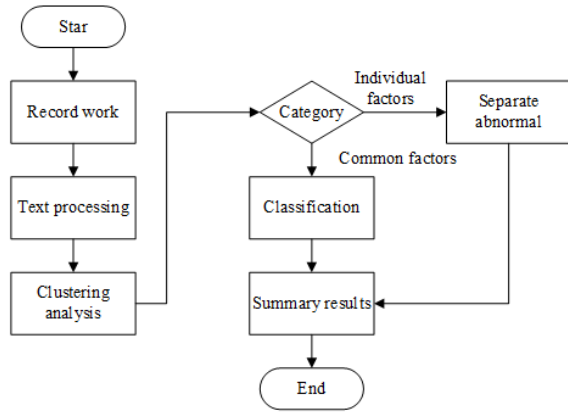
TABLE I. REALISTIC RECORD OF HEADLAMP INSTALLATION

Realistic table of site work			
Job name	**Body headlamp installation**	**Operator**	**Y&W**
Job describtion	People stand installing Body headlamp with operation tools, such as electric drilling machine, screwdriver		
No.	Job content	Total time	Remark
1	Rummage a suitable screwdriver from the toolbox	200s	1.many cars and materials are uniformly mixed in the middle of the workshop 2.The toolbox is located in the middle of the workshop
2	Get the electric rig and come back to the work area	200s	
3	Search lamp components that match the vehicle type from the material area	300s	
4	Take the headlight assembly (including the time to go back and forth when take the headlight assembly)	100s	
5	Remove the lamp assembly package	340s	
6	Squat down and search for a suitable plug from the material bucket (this bucket which contains various small parts is prepared by the workers)	100s	
7	Deal with the wiring on the car body and connect to the socket	200s	
8	Deal with the first headlight and connect the plug	240s	
9	Deal with the second headlight and connect the plug	250s	
10	Stand up and connect the first light to the car body and connect the two plugs	80s	
11	Adjust the position of the first lamp (aim at the mounting hole)	140s	
12	Squat down and search for suitable screws (three screws and three gaskets in total) from the material bucket	130s	
13	Stand up and fix three screws with a screwdriver	90s	
14	Stand up and connect the second light to the car body and connect the two plugs	70s	
15	Adjust the position of the second lamp (aim at the mounting hole)	160s	
16	Squat down and rummaged suitable screws (three screws and three gaskets in total, if there are improper screws, look for them elsewhere) from the material bucket	300s	
17	Stand up and fix three screws with a screwdriver	100s	
18	Adjust the position of two lampshades	120s	
19	Squat down and rummage for six screws from the material bucket	200s	
20	Stand up and fix two lampshades with a screwdriver	170s	
21	Pick up the tools and put them back in the toolbox	300s	
Total working time		3790s	

A. Realistic text processing

The text data should be processed by using data analysis before applying the K-means algorithm to process it. That means the natural language in the current realistic table should be processed into the language representation which can be recognized by the machine. It mainly includes word segmentation and vectorization of the text.

B. Text word segmentation processing

Chinese word segmentation processing is the basis of text data processing. Segmenting each word in a sentence and removing the text without practical meaning to get the standard text representation. Compared with the traditional word segmentation processing method, the word segmentation processing of the production site belongs to the field of professional processing. Therefore, how to establish a dictionary of term in the field of production is the key and difficult point to deal with. Such as "workers waiting for 20 minutes because of the abnormal material ", after the conventional segmentation processing for "due to /the reason/of/abnormal/material, workers/ need to/wait /20 minutes", adding in the field of professional knowledge dictionary word segmentation processing for "due to /abnormal material, workers /have to /prolong / 20 minutes' /unnecessary /working time", after remove the words without practical significance is expressed as "material defect leads to 20 minutes' waiting time of workers.

1) Establish a dictionary of workplace

Based on the traditional word segmentation dictionary, comprehensively considering the basic technology term, production management term, logistics term and the term expression that often appears in the site operation. Term expression on working site: "toolbox", "screwdriver", "screw", "nut", "take "," find", "rummage for", "binding", "distribution", "work area", "processing", "Squat down" ,"stand up", "walking time", "light components", "engine installation", "materials distribution", etc.

2) Establish a text dictionary for removing unnecessary words

This dictionary is aims to remove the meaningless words in sentences and standardize the expression of words. Invalid phrase Job content description, such as "the", "includes"," and so on ","yes", "and", "to", "by", "will", etc.

C. Vector expression of text

Before implementing clustering analysis, the text should be converted into a representation that the machine can understand. The commonly used text representation includes Boolean model, probability model and vector space model.

By use the vector model to represent the content description of the job. The vector space model expresses the text as a vector, as a point in the vector space, and expresses the similarity among semantics with the similarity in space.

1) Vectors expression

D represents the text set, t represents the feature item, and then the text is represented as D $(t_1, t_2\ldots, t_n)$. Generally, each feature item is given a certain weight to quantify. The vector with document D is expressed as D $(t_1, w_1; t_2, w_2; \ldots t, w_n)$.

2) Calculating weight

TF-IDF algorithm, the most effective implementation method of word weight and is adopted to calculate the weight value of the vector. IDF is called the inverse text frequency and is used to calculate the word's ability to distinguish text. Its calculation formula is as follows:

$$w_{ik} = tf_{ik}(t_k) * \text{idf}(t_k) = \frac{tf_{ik}(t_k)\log\left(\frac{N}{n_K} + 0.1\right)}{\sqrt{\sum_{k=1}^{n}\left\{[tf_{ik}(t_k)]^2 * \left[log^2\left(\frac{N}{n_k} + 0.1\right)\right]\right\}}}$$

D. K-means clustering analysis

K-means algorithm is an algorithm that inputs the number of clustering k and the database containing n data objects, then outputs the clustering that meets the minimum variance standard k. This topic uses k-means algorithm to realize the clustering analysis of text. The basic steps of cluster analysis are as follows:

The first step is to randomly select k objects from n data objects as the initial clustering center; the second step is to calculate the distance between each object and these central objects according to the mean value of each clustering object, and re-divide the corresponding object according to the minimum distance. The third step is to recalculate the mean value of each cluster. The fourth step is to calculate the standard measure function. When the function converges, the algorithm stops. If the condition is not satisfy, go back to step 2. The programming environment is Python + PycharmIDE + Scikit-learn machine learning framework. All the parameters of K - Means as follows, algorithm = 'auto', copy_x = True, init = 'K - means++, max_iter = 300, n_clusters = 4, n_init = 10, n_jobs = 1, precompute_distances =' auto ', random_state = None, tol = 0.0001, verbose = 0.(Set the value of n_clusters to 6, depending on the abnormal qualities in the workplace).

The first 50 analysis results are shown in Figure 2:

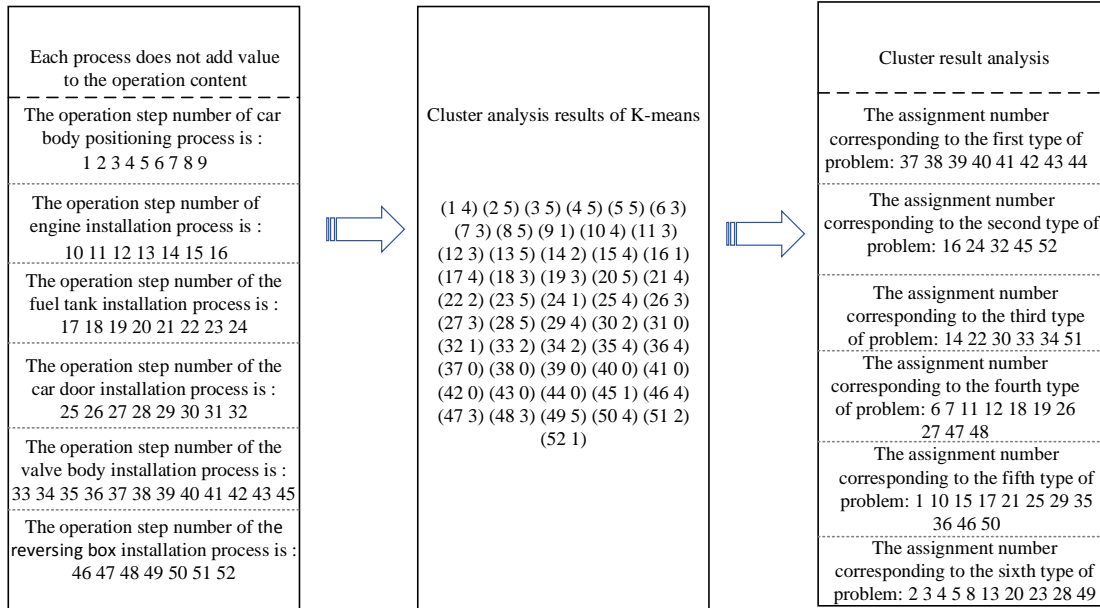| Each process does not add value to the operation content		
The operation step number of car body positioning process is : 1 2 3 4 5 6 7 8 9	Cluster analysis results of K-means	Cluster result analysis
The operation step number of engine installation process is : 10 11 12 13 14 15 16	(1 4) (2 5) (3 5) (4 5) (5 5) (6 3) (7 3) (8 5) (9 1) (10 4) (11 3) (12 3) (13 5) (14 2) (15 4) (16 1) (17 4) (18 3) (19 3) (20 5) (21 4) (22 2) (23 5) (24 1) (25 4) (26 3) (27 3) (28 5) (29 4) (30 2) (31 0) (32 1) (33 2) (34 2) (35 4) (36 4) (37 0) (38 0) (39 0) (40 0) (41 0) (42 0) (43 0) (44 0) (45 1) (46 4) (47 3) (48 3) (49 5) (50 4) (51 2) (52 1)	The assignment number corresponding to the first type of problem: 37 38 39 40 41 42 43 44

Figure 2. Cluster analysis result diagram

As shown in Fig.2, the cluster analysis result is processed as follows:

The main problems of the first type are: valve body installation associated with non-value-added operation;

The main problems of the second type are: the operation staff sorting out the packaging box and other garbage after the operation;

The main problems of the third type are: operating personnel rummaging in the material area for materials related to the operation;

The main problems of the fourth type are: the operator wastes a lot of time using crown-block crane to move the object to the station by experience;

The main problems of the fifth type are: the operator goes far away to get the tool;

The main problems of the sixth type are: the operator manually adjusts the heavy car body and carry the valve body according to the position of the crown-block crane.

IV CONCLUSION

Intelligent manufacturing has become the general trend of the manufacturing industry reform. The traditional manufacturing workshop continues to maintain relatively stable and efficient production, which is the urgent demand of enterprises. Highly manual production lines cannot operate in accordance with standardization, resulting in low overall production efficiency, which is also an urgent problem need to be solved.

Taking the M rail car manufacturing company as an example, this paper uses the K-means Algorithm of machine learning to dig out the common and individual problems existing in a wide range of production lines. Analyzing and classifying the non-value-added operations more rationally, saving time and cost. It provides more clear improvements in the design and layout of the assembly line.

REFERENCES

[1] Khosravi Roya,Hosseini Seyed Nezamedin,Javidanbardan Amin,Khatami Maryam,Kaghazian Hooman,Mousavi Nasab Seyed Dawood. Optimization of non-detergent treatment for enveloped virus inactivation using the Taguchi design of experimental methodology (DOE). [J]. Preparative biochemistry & biotechnology, 2019, 49(7).

[2] Wang Yanwei. Application and Optimization of Virtualization Technology in Metallurgical Rolling Production Line [J]. Chinese Metal Bulletin, 2018(10): 1-2.

[3] Huang Yuan. Introduction and application of k-means clustering algorithm [J]. Science Consulting (Technology Management), 2019 (10): 86-87.

[4] Li Feng, Li Mingxiang, Zhang Yujing. Fast K-means clustering algorithm for local iteration [J/OL]. Computer Engineering and Applications: 1-11[2019-10-21]. http://kns.cnki.net /kcms/detail/11. 2127. tp.20190815.1706.027.html.

[5] Jinxi Peng,Yuanqi Su,Xiaorong Xue,Donghong Song,Xiaoyong Xue. A parallel segmentation after classification algorithm of multi-spectral image of k-means of deep learning and panchromatic based on wavelet[P]. International Conference on Digital Image Processing,2019.