

Research on Oil Well Data Cleaning System

Yao Feng

School of Computer Science and Engineering
Xi'an Technological University
Xi'an, China
E-mail: 2352933419@qq.com

Li Zhao

School of Computer Science and Engineering
Xi'an Technological University
Xi'an, China
E-mail: 332099732@qq.com

Abstract—In the information age, with the continuous development of Internet technology, information data occupies every field of contemporary society. The development of the big data age makes these data more and more prominent. While users read the information they need from these massive data, data quality has also become a concern of users. A large number of data are preprocessed before data analysis, such as some duplicate values, missing values deal with inaccurate and other abnormal data, and filter the data through the data cleaning system to improve the standardization of the data, so as to improve the analysis efficiency of the data, reduce some unnecessary expenses, and save time and effort. The data cleaning system in this paper is implemented based on flash framework. Taking Python as the main language for data cleaning, technical cleaning and standard integration are carried out for some structural problems, duplication problems and missing problems of some different source data. Through the processing of abnormal data, the data quality and data analysis efficiency are greatly improved.

Keywords-Data Cleaning; Cleaning System; Oil Well; Outlier Handling

I. INTRODUCTION

Data is recorded with identifiable symbols or numbers to reflect objective things, such as pictures, numbers, words, etc. in the oil production

industry, the data value of each attribute reflects various indicators of oil wells, and each oil well has a certain range. Once it exceeds this range, it will affect the subsequent oil exploitation. In order to realize the normal exploitation of oil wells and realize the maximum oil production, data quality becomes very important. Therefore, users should clean some data affecting data quality from massive data, so as to improve the quality of data and make the oil exploitation work go smoothly. on this basis, this paper studies the processing of invalid data in oil wells. In order to optimize the data and facilitate the subsequent analysis, this paper proposes and designs an oil well data cleaning system based on Web.

II. RESEARCH BACKGROUND AND CURRENT SITUATION

In contemporary society, oil, as the backbone of modern industry, is of great significance to national development and is closely related to people's livelihood. Although China's oil and gas resources are relatively rich compared with other countries, compared with the world average, the per capital percentage of oil in China is still relatively low. Therefore, China must take effective measures to increase oil production and formulate reasonable plans to improve China's economy. Increase oil production and realize the maximum utilization of oil resources. Therefore,

the research on data cleaning system is very necessary for industrial production.

There are many researches on data cleaning at home and abroad [1], including: combining clustering algorithm, association rules and wavelet neural network to check and modify the abnormal data in the sensors and equipment that can be cleaned in the monitoring data. At the same time this method has a disadvantage that it needs to carry out secondary fitting correlation model for the data in the sliding window, [2] which not quickly clean the data that needs to be cleaned, In addition, the monitored data is regarded as the time series of each state quantity, and the iterative method is used to correct the abnormal data in the data. [3] This method also has an obvious disadvantage that there is a large deviation between the original value and the modified value, which causes some damage to the integrity of the data. There are also methods that can deal with outliers by constructing boundary models, but the processing results are inaccurate, which damages the continuity of data. There are also methods that use gravity search algorithm to identify errors and repair data [4], but the real-time performance of data is not very accurate. Therefore, the existing data cleaning methods only deal with the abnormalities of local state quantities, the processing of attribute correlation between data is not very good, and there is some damage to the continuity and integrity of data. Therefore, in order to reasonably clean the data, we should deeply express the data characteristics [5].

In recent years, with the rapid development of economy and the substantial improvement of information technology, many enterprises also have many problems in the design of data cleaning system, and there are still many aspects to be improved in the production of cleaning system [6]. Therefore, the manufacture of cleaning system has

a great application prospect for industrial production. It can depend on two factors:

- Compared with the original manual processing stage, the system can effectively process the abnormal data without spending a lot of time, manpower and financial resources to process these abnormal data, greatly improve the data quality and maximize the utilization of resources, which can not only save costs but also improve the utilization of resources.
- It has an immeasurable impact on China's economic development. The effective processing of data can obtain more accurate analysis results. According to these effective data, we can make appropriate adjustments to the mining scheme, improve oil production and be conducive to the country's economic development.

Based on the above two factors, we can understand that there is a lot of room to improve the development of oil well data cleaning system based on Web. The system is not only conducive to the effective utilization of resources, but also can save cost, time and labor, which has a great impact on national development and industrial production [7].

III. DEMAND ANALYSIS OF OIL DEPOT DATA CLEANING METHOD BASED ON WEB

As the starting point of system development, requirement analysis has guiding significance for system development, so requirement analysis plays a very important role. It is to make a comprehensive analysis of the problems to be solved, and then know what problems need to be solved. The person who makes the system must first understand what customers need, and then realize these in the developed system [8]. Demand analysis is the link between developers and

customers. Only by understanding the needs of users can we make a system to meet the needs of customers. Next, I will analyze the functional and nonfunctional requirements.

A. Functional Requirements Analysis

The system mainly provides a data cleaning platform for industrial production departments to clean oil well data. The functions of the system mainly include:

- Missing value processing

In the oil well data, there are many missing values due to mechanical and human reasons [9]. Mechanical factors are the failure caused by mechanical data collection or problems in saving data, such as memory damage, and human factors, such as arbitrary data modification and input errors. The existence of these values will have a great impact on the subsequent analysis results. Therefore, it is necessary to deal with these missing values, to improve data quality [10].

- Duplicate value processing

In the data set, there will be many duplicate values due to various errors. Some attributes can only appear once, so a large number of duplicate values will have an immeasurable impact on the data quality. Therefore, there is a great demand for the processing of duplicate values.

- Outlier handling

In the data, due to collection errors or system problems, some values that do not conform to common sense or values that do not conform to the defined data type may appear in the data. For example, abnormal data such as text and multiple decimals may appear in the data. In the process of cleaning [11], these abnormal values cause errors in the whole analysis and affect the whole progress of the analysis. Therefore, we must pay attention to the handling of outliers [12].

B. Nonfunctional Requirements Analysis

The evaluation of image quality can be divided According to the feasibility analysis, we can determine whether the production personnel will be satisfied with the cleaning system, according to the feasibility analysis of three aspects: technology, economy and operation. The data cleaning system is helpful for industrial production personnel to filter messy data, improve data quality and facilitate subsequent analysis [13].The following is a brief analysis of three specific aspects.

Technical feasibility: the system adopts the springboot framework and the main language is python. The system adopts the Android system architecture [14], and uses JSP technology to store information and standardize functional modules. Teachers and students can operate on mobile phones. According to the current mobile phone configuration, its own system performance is very high and it is very easy to realize.

Economic feasibility: the early data cleaning is based on manual operation. Under such operation, not only the data processing results are not ideal, but also a lot of time and energy will be spent on it. Before designing a data cleaning system, it is necessary for designers to think about its cost and the amount of resources. Catering to the needs of industrial producers, there will be objective economic benefits after development.

Operational feasibility: the system only needs to run on the computer, and then use the database to import the data to be cleaned and operate to process the data. The use process is not complex, the operation is simple, and it does not need to spend too much human and financial resources. It does not need professional skills to operate the system and process the data. It has strong interactive performance and good research significance, It is of great significance to industrial production [16].

IV. DESIGN OF OIL WELL DATA CLEANING SYSTEM

A. *Cleaning Principle*

In deep learning, the quality of data sets In the final analysis, data cleaning is to conduct in-depth analysis on the main causes of dirty data, find out the form in which these dirty data exist in a large number of data to affect the analysis results, use the current existing technology and optimize its processing to find dirty data, and change these detected data into values that meet normal needs after processing. The idea of data cleaning mainly uses the backtracking method [17], which takes the detected data as the starting point, makes a systematic analysis of each detail of the data flow direction, and summarizes a widely used data cleaning algorithm and related rules, which can be applied to various types of data cleaning systems [18].

B. *Web Implementation*

The web was mainly used for browsing static interfaces in the early days. These static interfaces are written in HTML and placed on the server. Users use the browser to request the web page on the server through the HTTP protocol. After the [19] web server software obtains the request sent by the user, it reads the identified resources [15]. When the client receives the browser sent by the message header, the browser parses the HTML data of the response, Show different vivid HTML pages to the client. However, with the continuous development of the network, more and more businesses begin to develop online, so the web application based on the network has become not simple. The resources accessed by users are not only the static pages stored on the server hard disk. A large number of applications require the web page information dynamically generated based on the user's request. It is even more difficult to extract some relevant information from the database and return a

generated page to the customer through some operations.

How to achieve:

It is realized through the server software of HTTP protocol, which reserves the interface that can be extended in advance [20]. What the user needs to do is to provide relevant expandable functions according to the rules. When the client request is transmitted to the web server software, it will judge whether it is to visit the relevant extended functions provided by the user. If it is judged to be yes, the relevant programs written by the user will be processed. After the processing is completed, the program will return the processing results to the client.

C. *System Framework*

Flask is a web framework of python, which is very popular at present. The flexibility of Python language also makes flask have the same characteristics. Therefore, the biggest advantage of flask is that it is simple and lightweight, and developers can flexibly use the features to be developed, Let developers freely and flexibly compatible with the features to be developed. Whether it is user portraits or product recommendations, python has a very heavyweight advantage. Because flash is relatively light and simple, it is easy to start, and the cost of trial and error is low. Flask's framework also brings a more image extraction mechanism, which can work on almost any platform of Python. Flask's data access framework can also solve some difficulties that occur frequently when using the database [21].

D. *Function Specific Design*

- Pretreatment stage

Preprocessing is divided into two steps: one is to adjust the data to an uploadable file. Second, look at the data. Through manual viewing, you can have a more intuitive understanding of the data,

find some obvious errors at the beginning, and make some simple preparations for later processing.

- Missing value processing

Missing value is the most common problem in data. There are many filling methods for missing value: one is to fill in the missing value according to their own knowledge or experience; the other is to fill in the missing value according to the calculation results of the same indicators, such as filling in the missing value according to the calculation results of different indicators; the third is to fill in the missing value according to the calculation results of different indicators. The mean filling method is adopted in this design. The calculation results of the same index are used for summation, and then the average value is taken to fill in the missing data. This method has the advantages of simple calculation, fast filling speed and close to the real value [22].

- Duplicate value processing

In some attributes, some data can only appear once, and sometimes errors will occur, resulting in data duplication. Therefore, the processing of duplicate values should be realized.

For the processing of duplicate values, first find the duplicate values in the data set and save them in one list, save the duplicate index to another list, and then take out the average value of the current attribute and take out the index value for replacement.

- Outliers

Some unreasonable values will appear in the data, such as text classes or abnormal values beyond the normal range in the data set.

For values beyond the range, we will check and clear them according to the previously defined range, and then use the filling method. For text and

decimal types, we can clear them by defining data types, and finally use the missing value filling method.

E. Flow chart system function

According to the brief description of specific functions, the relevant functional flow chart of oil well data cleaning system can be obtained. The functional flow of the system is shown in Figure 1.

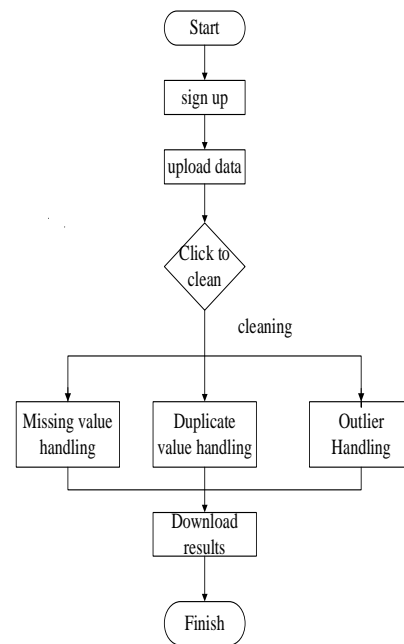


Figure 1. System function flow

F. Database implementation

MySQL is a relational database, which can be used in a variety of storage because it has multiple storage engines. It has many advantages: high speed, small volume, easy installation, etc. Most developers use it as the primary database. Now, under the name of Oracle, it is a kind of well used and free database, which can be used in small and medium-sized databases. If it is designed to be easy to use and support all kinds of locks, it can be used well in concurrent processing, and the types of MySQL fields meet the needs of most designers. The addition of different types of transactions also

makes the storage of data more favorable and improves the accuracy. It is not easy to cause repeated write and other problems. The management of permissions makes the distinction between superiors and subordinates more perfect, and it becomes easier to set permissions, so that the security of data is more and more guaranteed. Based on the above excellent points, I take MySQL database as the main database I develop and use. Its driver management is shown in Figure 2:

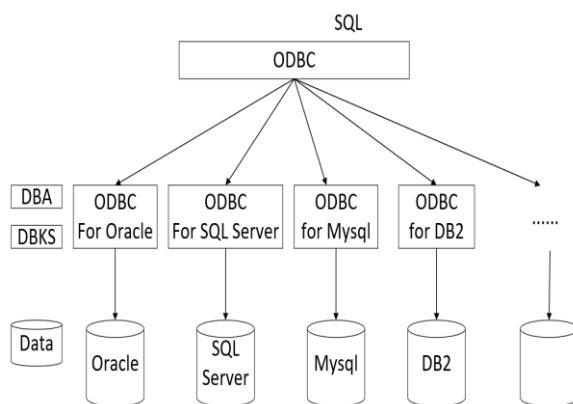


Figure 2. Database driver management diagram

V. DATA CLEANING SYSTEM OF OIL DEPOT BASED ON WEB

A. Function description

The web-based oil well data cleaning system is divided into two parts: the front-end interface display and the background data processing system, which are implemented based on the flash framework. The specific functions are as follows:

User: the user registers and logs in. After the user logs in successfully, the page jumps to the home page [25]. The user uploads the relevant oil well data to be cleaned according to the needs. After parsing the excel file, the data cleaning system cleans the data in the file. After cleaning, a link will be returned. You can apply to download the cleaned results by pasting it into the browser.

Specific cleaning rules need to be defined by ourselves.

Its specific functions are as follows:

- User login registration

When users log in the system for the first time, they need to register to determine the login account. Users need to fill in their own user name and set the login password. When logging in directly, the system will automatically prompt whether there is a login account and whether the password is wrong. If there is an error, the system will enter it for the second time.

- User information management

Users can modify their own information in my home page. The information filled in when logging in for the first time is saved on my home page.

- Data cleaning

The user uploads the data to be cleaned, and then the system automatically parses it. After parsing, the data is processed according to the set range, mainly dealing with some missing values, duplicate values and abnormal values for embedding function, and can also use the tag of JSP to parse and convert it into HTML tag, or use XML tag.

- Download Results

After that, you can choose to download and save the cleaned data locally, and users can conduct in-depth analysis. Facilitate the next step of work.

B. Cleaning data upload

After data cleaning, a link will be returned. Copy the link to the browser to view the cleaned results, download and save the results. As shown in Figure 3.



Figure 3. Data import

C. Cleaning result output

The collected data is shown in Figure 4. The results after cleaning are shown in Figure 5. According to the comparison between figure 4 and figure 5, we can see that the red box indicates that the system fills in the missing value, the green box indicates that the missing value is filled in after the duplicate value is also deleted, and the yellow box indicates that the abnormal value is deleted and filled in. The system has basically realized the cleaning of messy data.

2	35-55	3.5				2.5		11.67	73.6	3		
3	54-61	5.3	11	229.75		5.4	13.97	51.6	4	1.9454		
4	21-51	3.4	32.93			2.42	8.37	59.88	3			
5	56-46	2.8					6.96	57.7	2		2.82	
6	31-49	3.9	38.17	244.83	2.49	3.24	11.75	71.13	3			
7	32-65	3.1	22.37	262.56	2.36	3.34	15.23	70.35	2	15.6	2.4	
8	32-65	7.3	15.02	239.97	2.44	3.75	10.79	53.8	2	25.1	3.5	
9	29-54	2.8	46.94	235.14	2.51	3.61	9.66	61.4	2			
10	29-54	6	15.6	251.31	2.45	2.78	12.19	50.8	3			
11	34-28	5.5	48.42	210.75	2.55	6.8	5.65	55.8	3	18.0+0.5	2.2*2.3	
12	34-57	5.4	11.74	243.55	2.52	12.24	11.5	46.05	3	30.0+0.6	2.4-2.5	
13	34-54	2.8	26.54	239.97	2.47	8.4	10.5	60.3	3	25.0+0.51	2.04	
14	34-54	4.1	19.36	232.11	2.51	12.04	9.2	51.41				
15	13-67	3.8	17.9	231.6	2.55	9.9	9.1	52.9	3	14.5	1.7	
16	13-67	1.6	34.8	224.4	2.52	5.4	7.9	60.1				
17	05-105	4	38.46	229.75	2.5	5.8	8.8	62.45	2	26.7	2.82	
18	05-105	3.9	18.23	251.23	2.43	7.1	12.37	57.67	3			
19	32-23	1.4	17.7	238.3	2.51	4.2	10.47	49.3	0.2	15	3.5	
20	13-65	5.3	16.06	243.13	2.49	9.47	11.03	52.4	3	25.6	1.95	
21	34-19	6.9	31.57	213.57	2.62	17.59	5.47	13.3	0.2	23.8	3	

Figure 4. Data collection

2	35-55	3.5	31.37	234.57	2.5	6.91	11.67	54.38	3	17.65	3.1	
3	54-61	5.3	31.07	229.75	3.06	5.4	13.97	51.6	4	1.9454	3.52	
4	21-51	3.4	32.93	234.23	2.42	7.23	8.37	59.88	3	17.29	3.43	
5	56-46	2.8	30.93	234.49	3.22	6.64	10.89	57.7	2	16.89	2.82	
6	31-49	3.9	38.17	244.83	2.49	3.24	11.75	54.6	3	17.64	3.63	
7	32-65	3.1	22.37	234.11	2.36	3.34	11.47	54.14	2	15.6	2.4	
8	32-65	7.3	31	239.97	2.44	3.75	10.79	53.8	2	25.1	3.5	
9	29-54	2.8	46.94	235.14	3.25	3.61	9.66	61.4	2	17.1	3.15	
10	29-54	6	31.45	234.8	2.45	2.78	12.19	50.8	3	17.1	3.71	
11	34-28	5.5	48.42	210.75	3.36	6.8	11.11	55.8	3	17.48	2.96	
12	34-57	5.4	31.48	243.55	3.01	12.24	11.5	46.05	3	17.84	3.47	
13	34-54	2.8	26.54	239.97	2.47	8.4	10.5	60.3	3	17	2.04	
14	34-54	4.1	31.77	232.11	3.07	12.04	9.2	51.41	4.19	16.99	3.61	
15	13-67	3.8	31.1	231.6	2.78	9.9	9.1	52.9	3	14.5	1.7	
16	13-67	5.39	34.8	224.4	3.39	5.4	10.84	60.1	4	17.6	3.04	
17	05-105	4	38.46	229.75	2.5	5.8	8.8	62.45	2	26.7	2.82	
18	05-105	3.9	31.86	234.58	2.43	7.1	12.37	57.67	3	17.37	2.87	
19	32-23	5.35	31.73	238.3	3.39	4.2	10.47	49.3	4.16	15	3.5	
20	13-65	5.3	31.45	243.13	2.49	9.47	11.03	52.4	3	25.6	1.95	
21	34-19	6.9	31.57	213.57	2.84	6.98	11.55	13.3	3.57	23.8	3	

Figure 5. Cleaned data

D. Functional testing

The results are shown in Table 1. According to the above tests, we can know that the pre designed functions of the system can be basically realized smoothly without other errors during operation.

E. Operating environment

JSP is a popular service site at present. You can view it when you need JSP interface and input the website you want to place them. JSP is a back-end language. It is developed based on Java code. It shows a dynamic effect by inserting the data obtained from the operation of Java code in the file into the static page. It can use the tag of XML file Using this configuration can reduce the dependence of the code on the server platform and run everywhere, which is also due to the Java virtual machine. JSP compiler can convert the code into the original binary code, and then run the binary code directly, which is one of the advantages of JSP. Compared with pure servlets, JSP can easily write or modify HTML web pages without facing a large number of print statements. It reduces the burden of development, so I use JSP as my page language.

TABLE I. TEST CONTENTS AND RESULTS

Test content	Expected results	Actual results	judge
Enter user name and password	Login successful	Same as expected	adopt
Upload correct data sheet	Upload successful	Same as expected	adopt
Upload error data table	report errors	Same as expected	adopt
Are duplicate values processed	Processed	Same as expected	adopt
Are missing values populated	Filled	Same as expected	adopt
Are outliers handled	Processed	Same as expected	adopt
Whether data can be saved	Successfully saved	Same as expected	adopt

VI. SUMMARY

The oil well data cleaning system based on Web has great potential and development value. The implementation based on web in this paper integrates the design mode, further improves the scalability and maintainability of the code, improves the concurrency of the system, and improves the concurrency ability of the system. When a large amount of data needs to be cleaned, it can ensure the normal operation of the system. Data mining technology is widely used in medical treatment, energy, retail, automobile, finance and many other fields to provide decision-making and suggestions by mining valuable information. Medical data mining can provide targeted guidance to patients, predict the change and development trend of physical health status, and take preventive measures. Accurate analysis can reduce over treatment and insufficient treatment; Use big data to analyze energy purchases, so as to predict energy consumption, improve energy efficiency and reduce costs by managing energy users; For retail enterprises, data mining technology can well integrate all kinds of information, help enterprises master customer

needs, and realize precision marketing and personalized services; With the help of data mining technology, insurance companies can fully understand drivers' driving habits and driving behavior, provide different types of insurance products, and so on. The widespread existence of dirty data leads to the limited available data in the process of data mining. Data cleaning is particularly important. Different data cleaning methods have their own advantages and disadvantages, and the research on data cleaning methods will be more in-depth.

REFERENCES

- [1] Peng Nan Research and implementation of cleaning rule mining system for dirty data [D] Shanghai: Donghua University, 2018.
- [2] Hu daiguo Design and implementation of data cleaning scheme for mobile comprehensive information system [D] Shandong: Shandong Normal University, 2018.
- [3] Qu Wanrong Urban road traffic flow data cleaning technology and system implementation [D] Zhejiang: Zhejiang University of technology, 2017.
- [4] Wang Jiang Research on data cleaning technology and design and implementation of cleaning framework [D] Inner Mongolia: Inner Mongolia University, 2016.
- [5] Wang Kui Research and application of Hadoop based traffic data cleaning [D] Guangdong: Guangdong University of technology, 2017.
- [6] Sun Chang Research on SCADA data cleaning method of wind turbine [D] North China Electric Power

- University; North China Electric Power University (Beijing), 2018.
- [7] Li Ningning Research and implementation of optimization technology in big data cleaning system [D] Heilongjiang: Harbin Institute of technology, 2016.
- [8] Wang Liping Research on on-line cleaning technology of heat exchanger based on scaling and scale dissolution mechanism [D] Beijing: Beijing University of chemical technology, 2017.
- [9] Zhang Xiaoli Research on data cleaning framework and missing value reconstruction method [D] Shaanxi: Xi'an University of technology, 2015.
- [10] Fang Xiaoli, Liu Xia Research on data cleaning method for document visual analysis [J] Journal of University Library and information science, 2021,39 (06): 56-60.
- [11] Yuan Yanwei, Xu Ling, Ji Fuhua, Guo Dafang, an SA, Niu Kang Big data cleaning method and experimental optimization of agricultural machinery operation [J] Journal of agricultural machinery, 2021,52 (06): 35-42.
- [12] Xu Sijia, Wang Xiang, Zhao Chenglin, Xu Fangmin Research on data cleaning methods of industrial machinery and equipment [C] // . Proceedings of 2020 China information and communication conference (CICC 20202020:347-352.)
- [13] Data Cleaning in Cloud Platform [J]. International Journal of Recent Technology and Engineering, 2020, 9(1).
- [14] Jinlin Wang, Xing Wang, Yuchen Yang, Hongli Zhang, Binxing Fang. A Review of Data Cleaning Methods for Web Information System [J]. CMC: Techniques Computers, Materials & Continua, 2020, 62(3).
- [15] Yogita Bansal, Ankita Chopra. Data Cleaning for Large Data Sets [J]. International Journal of Recent Technology and Engineering (IJRTE), 2020, 8(6).
- [16] Yanli Bai. Data cleansing method of talent management data in wireless sensor network based on data mining technology [J]. EURASIP Journal on Wireless Communications and Networking, 2019, 2019(1).
- [17] Candelotto Laura, Grethen Klara J., Montalcini Camille M., Toscano Michael J., Gómez Yamenah. Tracking performance in poultry is affected by data cleaning method and housing system [J]. Applied Animal Behaviour Science, 2022(prepublish).
- [18] Liu Shengjie, Li Guangye, Jiang Shize, Wu Xiaolong, Hu Jie, Zhang Dingguo, Chen Liang. Investigating Data Cleaning Methods to Improve Performance of Brain-Computer Interfaces Based on Stereo Electroencephalography [J]. Frontiers in Neuroscience, 2021, 15.
- [19] Shi Xi, Prins Charlotte, Van Pottelbergh Gijs, Mamouris Pavlos, Vaes Bert, De Moor Bart. An automated data cleaning method for Electronic Health Records by incorporating clinical knowledge [J]. BMC Medical Informatics and Decision Making, 2021, 21(1).
- [20] Àvila Callau Aitor, PérezAlbert Yolanda, Serrano Giné David. Quality of GNSS Traces from VGI: A Data Cleaning Method Based on Activity Type and User Experience [J]. ISPRS International Journal of Geo-Information, 2020, 9(12).
- [21] Feiyu Lian, Maixia Fu, Xingang Ju. An Improvement of Data Cleaning Method for Grain Big Data Processing Using Task Merging [J]. Journal of Computer and Communications, 2020, 08(03).
- [22] Information Technology - Data Systems; Studies from Harbin Institute of Technology Reveal New Findings on Data Systems (A Review of Data Cleaning Methods for Web Information System) [J]. Information Technology, 2020.
- [23] Domenico Vitale, Gerardo Fratini, Massimo Bilancia, Giacomo Nicolini, Simone Sabbatini, Dario Papale. A robust data cleaning procedure for eddy covariance flux measurements [J]. Bio geosciences, 2020, 17(6).
- [24] Jinlin Wang, Xing Wang, Yuchen Yang, Hongli Zhang, Binxing Fang. A Review of Data Cleaning Methods for Web Information System [J]. CMC: Computers, Materials & Continua, 2020, 62(3).
- [25] Science - Analytical Science; Researchers from Southeast University Discuss Findings in Analytical Science (Robust data cleaning methodology using online support vector regression for ultra-short baseline positioning system) [J]. Journal of Engineering, 2020.