*Research Article*

# Application of Data Mining in Performance Management of Public Hospitals

## Huiqun Lu [ID],[1] RuiLing Wang,[1] and Zhenju Huang[2]

[1]*Personnel Department, Nantong First People's Hospital, Nantong, Jiangsu 226001, China*
[2]*High Quality Development Management Division, Nantong First People's Hospital, Nantong, Jiangsu 226001, China*

Correspondence should be addressed to Huiqun Lu; xdrc88@ntu.edu.cn

With the rapid development of computer technology, information technology covers all aspects of daily life, and the medical industry is also paying more attention to information construction. Conventional management methods have been unable to further improve the hospital's management capabilities. At the same time, countries that are better in terms of hospital management practices have set a benchmark for mainland hospitals and reformed hospitals in order to stand out in the future. In addition to evaluating the economic benefits and work efficiency of doctors, hospitals must also consider that hospitals, as a special service industry, cannot be measured by economic indicators. Therefore, there is a multiparty game in the performance appraisal of hospitals, and it is necessary to consider not only economic factors but also the characteristics of public services. This article is based on the case of a large domestic tertiary hospital, combined with the hospital's performance management reform plan, through the design idea of performance management and incentive performance pay distribution, using data mining technology as an auxiliary means. It successfully helped the hospital complete the performance and incentive performance pay aspects reform. The main research work of this paper is divided into the following three aspects. (1) Using data mining technology, according to each nursing unit's workload, risk level, the difficulty of internship, and other objective factors in the past year for patient outpatient visits, surgery implementation, critical first aid, etc., are classified in line with the actual situation and provide a reliable basis for the reasonable and efficient allocation of hospital human resources. (2) In the performance management system, we integrate the third-party data mining tool weka to assist in the evaluation of the performance distribution plan and the calculation of the follow-up incentive performance pay. (3) We use the mathematical model of data mining to measure and evaluate the reasonableness of historical workload and performance appraisal, determine a new incentive performance pay distribution model, and use the software as a calculation tool for the internal distribution of performance wages to provide monthly incentive performance wage statistics in the future.

## 1. Introduction

The opinions of the Central Committee of the Communist Party of China and the State Council on Deepening the Reform of the Medical and Health System issued by the Ministry of Health mentioned that the overall goal of the new medical reform is to establish and improve a basic medical and health system covering urban and rural residents and provide the people with safe, effective, convenient, and inexpensive health services, which is what we often say to solve the difficult and expensive medical treatment. According to the preliminary budget of the Ministry of Finance, in three years, governments at all levels need to invest a total of 850 billion yuan to support the five key reforms of the basic medical security system. The resulting investment in information construction will reach 17.05 billion [1–8]. Among the major fields that support the development of hospitals, medical informatization has a unique perspective and has become the only technical pillar among the four pillars and eight pillars. With the gradual expansion of the scale of hospitals, the difficulty of management is increasing. Taking into account the economic and social responsibilities of hospitals has become a difficult problem for every dean. The performance management of

hospitals is the same as that of enterprises. Through the establishment of an incentive system, employees are effectively encouraged to exert their subjective initiative and improve work efficiency. However, the operation of a hospital cannot be calculated simply by money, and the success of a hospital cannot be measured solely by income. To perform performance management in a hospital, it is necessary to consider not only the incentive mechanism to promote the work efficiency of medical workers but also the necessity of certain special positions, so that employees cannot feel the impact of the growth of the hospital on themselves, nor can they allow this influence occupies absolute dominance and makes the hospital become profit-oriented. How to weigh the feasibility of the performance management plan has also become one of the most complicated topics in the current hospital management [9–16]. Only by quantifying the workload of medical staff as much as possible, objectively evaluating the labor output and benefits of relevant personnel, linking the evaluation results with team and individual incentive performance pay, and establishing a corresponding reward and punishment mechanism can the system be promoted. Medical workers complete the medium and long-term goals set by the hospital and finally realize the hospital's strategic plan. On the other hand, since performance reform is directly related to the vital interests of employees, it is necessary to be cautious in the reform process. It is necessary to touch the interests of some late-entry departments but also to maintain a certain balance; it is necessary to increase the economic income of certain high-tech departments and to take into account the enthusiasm of the general staff at the grassroots level. Therefore, performance reform should not cause excessive internal turbulence. It is necessary to maintain the status quo of hospital operations to a certain extent but at the same time change appropriately and seek a balance between changed and unchanged, dynamic, and immobile. Therefore, as a provider of a performance reform program, it must include a complete set of mechanisms. Through management consulting, strategic planning, measurement, and evaluation, the principle of using data to speak should be implemented from beginning to end to ensure fairness. The content of this paper is based on a successful case. Through the use of data mining technology, the historical data of a certain hospital is processed and analyzed, and the mathematical model is used to analyze the performance appraisal plan, and then the performance appraisal results are calculated periodically through the software. It gives incentive performance pay to various departments. Combining computer software with data mining technology to realize performance appraisal program analysis and incentive performance pay distribution, this article is based on the implementation of a set of self-developed performance management software in a domestic large-scale tertiary hospital as an example.

The rest of the paper is organized as follows. Section 2 focuses on related work, while Section 3 throws light on performance management systems based on data mining. Similarly, Section 4 is about evaluation and discussion and the final section is the concluding section.

## 2. Related Work

The concept of performance management was first proposed by American management scientist Aubrey Daniels in the 1970s. In 1842, the U.S. federal government implemented performance evaluation on civil servants [17]. The U.S. federal government has conducted comprehensive performance evaluations on thousands of federal projects [18]. Performance management is a complete system, usually composed of measurement, evaluation, and management sequence. In this system, employee performance is planned in advance and supervised, instructed, and evaluated in the process, which is finally achieved by improving personal performance corporate strategic goals [19]. There are similarities between the performance of medical institutions and the performance management of enterprises, and they often learn from the methods of corporate performance management for performance management and performance appraisal. We can consider performance as an institution, the results of the behaviors, activities, procedures, and actions of its members, and institutions and members expect this result to be achieved with the least resources [20]. Previously, hospital performance management mainly refers to KPI (Key Performance Indicator), BSD (Balanced Score Card), and other methods in affairs [21]. The BSD method formulates management objectives from four aspects, namely, finance, customers, internal operations, and learning growth. Corresponding evaluation indicators are set for different specialties and different levels. The original intention of KPI is to encourage owners, contractors, suppliers, and other engineering project participants to accurately evaluate their own performance [22]. The U.S. currently generally uses the RBRVS (resource-based relative value scale) evaluation system to evaluate the performance of clinical medical staff. RBRVS is a research obtained by Harvard University in the United States. It is mainly used to objectively measure the resources consumed by physicians in providing services to patients [23]. Domestic hospitals generally adopt two methods for performance appraisal: income and expenditure balance and KPI. The former is more intuitive. By subtracting expenditures from income and multiplying coefficients, the economic benefits are clear at a glance. The latter assesses medical workers by formulating a small number of indicators, and the operation is relatively simple. However, the former is likely to cause hospitals to ignore social responsibilities, and the latter is often misunderstood as the finer the better. In order to solve these problems, hospitals are headed by many tertiaries [24, 25].

One of the important topics in the field of database knowledge is knowledge discovery (KDD). The term KDD originated from the 11th International Conference on Artificial Intelligence held in Detroit, USA, in August 1989. At the first KDD&DM international academic conference held in Montreal, Canada, in 1995, data mining technology was divided into knowledge discovery in the field of scientific research and knowledge discovery in the field of engineering [26]. After that, such a conference is held once a year. After years of hard work, the research on the neighborhood of data

mining technology has made great progress. The newer developments in the world are mainly manifested in further research on the methods of discovering knowledge. For example, in recent years, attention has been paid to the research and improvement of the Bayesian method and Boosting method, the close integration of KDD and database, and the application of traditional statistical regression methods in KDD. In terms of application, it is mainly reflected in the KDD commercial software tool from the isolated process of problem-solving to the establishment of an overall solving system. The main users are insurance companies, large banks, and the sales industries. Many computer companies and research institutions attach great importance to the development and application of data mining. Both IBM and Microsoft have established corresponding research centers. The United States is the most prosperous region for data mining research in the world, and it occupies the core position of research [27]. Domestic data mining started relatively late, and it is currently in the development stage. For example, in the research neighborhood of classification technology, it is trying to establish its set theory system while realizing the processing of massive data; in the process of KDD, the rough set and fuzzy set theory are used comprehensively; the fuzzy system knowledge model and fuzzy system are constructed. There are identification methods, related research on theoretical modeling and technical realization of Chinese text mining, and text mining using concepts [28]. Data mining technology is currently used in a wide range of fields. It can be seen in various fields such as finance, telecommunications, rapid consumption, e-shopping, biomedicine, and social networks. Data mining technology is used to further improve the company's industry competitiveness, through knowledge discovery, to obtain previously unavailable or unknown information. Once this information is used in the industry, it can often help companies make more profits and provide more data for scientific research.

## 3. Performance Management System Based on Data Mining

*3.1. Business Analysis.* In the entire program and system construction, the most difficult and most important thing is to realize the transformation of the incentive performance pay accounting method, from the income-based method to the workload-based evaluation method. At this point, the hospital has high requirements for the software. After analysis, this part of the business mainly includes four parts: model establishment and maintenance, model measurement and evaluation, incentive performance pay calculation, and result query. Among them, the establishment and maintenance of the model and the calculation and evaluation of the model are the most complicated, which is also the focus of this article.

(1) Model establishment and maintenance: it is necessary to consider the classification of hospital employees. The in-service employees of the hospital mainly include doctors, nurses, medical technology,

and management and logistics. The nature of the work of the four types of personnel is completely different. Starting from the workload, this assessment focus of the four categories of personnel is also different.

(2) Simulation calculation and evaluation: one of the important missions of this performance management is to change the method of performance appraisal, which is to use the new performance model to assess the staff of the whole hospital. Then, whether the new model is established and whether the coefficients of the model are reasonable must be tested by actual data and finally approved by the staff congress of the whole hospital to establish. In order to facilitate the evaluation and find the problems as early as possible, it is necessary to apply the new model to the historical data and generate the historical data in the same period. Comparing the actual incentive performance pay that occurred last year with the current simulation data over the same period, after evaluation and judgment by relevant departments, the incentive performance pay model is confirmed after the results are recognized. Otherwise, the model will be adjusted, and the evaluation will be recalculated.

(3) Calculation of incentive performance pay: after the model is established, the system should automatically calculate incentive performance pay for each department on a monthly basis since the project is officially launched, and the background should also be able to automatically obtain data from each system. At the same time, some data cannot be automatically obtained through the existing system, and the system is required to support the manual filling.

(4) Result query: after the performance appraisal is over, users can independently check the final result of their appraisal, including incentive performance pay. At the same time, it is also possible to obtain the complete original data of its performance appraisal through detailed reports, so as to facilitate its verification.

*3.2. System Requirements Analysis.* The role of the performance management system is to assist performance reform, by changing the traditional performance evaluation methods, forming a new workload-oriented performance evaluation system. The system needs to play a role in linking up and down; that is, it is compatible with the assessment plan that has been used for many years. It will not continue to carry out performance reforms because some departments are temporarily unable to pass the workload assessment, adopting new accounting methods, fully embodying the core idea of more work and more rewards, creating clear guidance for hospital staff, and guiding staff to make more efforts in technology, scientific research, and difficult and critical directions.

On this basis, the main functional requirements of performance management can be divided into the following parts. First, the need to establish a data warehouse must meet the effectiveness of performance appraisal, record the actual data that occurs as of the end of each month, and, at the same time, can also improve the efficiency of incentive performance pay calculation. Second, the demand for modeling and measurement requires not only the creation of targeted incentive performance pay models for different job categories but also the traditional performance appraisal methods. The system should generate simulation data as soon as possible, provide evaluation and analysis, and shorten the evaluation cycle. Third, concerning performance evaluation and incentive performance pay calculation requirements, once the model is determined, the system should automatically calculate the performance evaluation results and incentive performance pay of each accounting cycle in the future. Fourth, various types of reports for users' inquiries are provided. Fifth, different levels of security mechanisms are provided.

### 3.3. Data Warehouse Design.

For the usual performance management system, there is no such design as a data warehouse. Whether it is a management method based on KPI (key performance indicators) or performance management based on BSD (balanced scorecard), there is no use of historical data for modeling or evaluation. The requirements and design plan of this performance management system are different from the past. In order to carry out performance reform in a hospital that has implemented a performance plan for many years, it is necessary to take both reform and historical status into consideration. How to determine the new incentive performance pay base of each department and how to reflect the orientation of the hospital needs to be answered through historical data analysis. In some future periods, similar problems may appear again. The performance management system designed in this article is actually to use data to solve the hospital's confusion.

### 3.3.1. Data Warehouse Architecture.

When the informatization of the medical industry reaches a certain level, various information systems will inevitably appear. In this project, the in-depth analysis should be done through the data that already exists in these systems to help the management make decisions.

This project adopts the data warehouse method to provide data sources for the performance management system and adopts the ETL method to provide the data warehouse with sorted and cleaned data. Taking into account the characteristics of hospital performance management, the designed data warehouse covers finance, cost, human resources, workload, attendance, scientific research, and many other fields, and it also combines questionnaire surveys to ensure the integrity of information. Figure 1 is the architecture diagram of the data warehouse.

At the same time, considering that performance management is a cyclical work, building on a data warehouse can better reflect the effectiveness of its data. Based on these
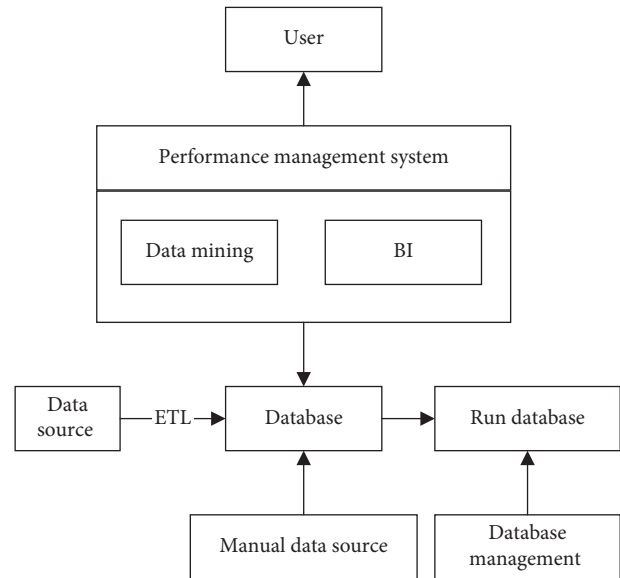


Figure 1: Architecture of data warehouse for performance system.

analyses, the requirements for building a data warehouse to achieve data management have been clarified. To build a data warehouse, first, we perform modeling according to business classification and determine the data structure provided to the upper application system. Then, according to the required business model, we perform ETL of data from each information system and load the data into the business model. ETL is divided into two layers. Firstly, the data is loaded from the data source to the data warehouse.

The structure of the data is kept as it is to ensure the efficiency of data collection; then, it is handed over to the data warehouse for ETL to organize the data into a business model and store in the required format.

### 3.3.2. Table Structure Design of Data Warehouse.

In order to meet the application of data mining, the data warehouse should include tables and their structures as follows:

(1) HIS execution data: the main business process of the hospital relies on the HIS system; that is to say, the HIS system is the core system that records the workload of the hospital's clinical frontline staff. Due to its positioning problem, the HIS system focuses more on charge management. In the performance appraisal plan, the charge schedule can be used as a statistical table for department billing and execution projects.

(2) The incentive performance pay data of the department: the incentive performance pay data of the department comes from the financial or tax reporting system, which records the monthly incentive performance pay income of the whole hospital. As the basic data for performance appraisal and calculation, the accounting department needs to provide department incentive performance pay that is accurate to the department and separate medical

and nursing care. The hospital's incentive performance pay data is accurate to the employee, with the employee's name and employee number. Unfortunately, there is no information on the department mentioned by the employee at that time. After coordination, the hospital reformed the tax reporting system and improved the department information.

(3) Cost data of the department: cost data is collected from the cost accounting system, and the cost items are classified in a detailed manner, including employee wages, house leasing, water, electricity and coal, disinfection supply, night shift fees, catering expenses, and travel expenses. Performance appraisal only uses the controllable cost part.

(4) Special data: obtaining data from existing systems through computer technology is a convenient and quick method, which can greatly reduce manpower and physics. However, sometimes, certain businesses are not supported by a fixed system, such as simple questionnaires, which are distributed to customers through paper media, or return visits to medical quality by telephone. Such surveys are difficult to formulate corresponding surveys. The system and the operability of the system's data entry also have problems. For such special data, there is a way for users to fill in the data manually in the data warehouse, and all the filling contents are simply designed in column format.

(5) Surgical records: surgical records are collected from the hospital's surgical record system. The main function of this system is to count the name of the operation, participants, consumables, etc. The operation records required by the performance system are the operating department, the name of the operation, and the time. Because it does not involve the secondary distribution of incentive performance pay, there is no need for specific personnel information.

(6) Department dictionary and project dictionary of HIS: this type of dictionary data is relatively fixed, and its main function is to assist the division of HIS departments and the attribution of items.

### 3.3.3. Build Data Warehouse.

In addition to the HIS system in the hospital, the systems already in use at the same time include the financial system, cost accounting system, and human resource system. There is no data synchronization between the system and the other system. The same person's information is expressed in several systems from their own perspectives. To build a data warehouse, you first need to solve the problem of information islands. After analysis, it is found that the two types of information that are most likely to cause conflicts in these systems are personnel and departments.

As long as there are the establishment of unified personnel and department standards, the establishment of the mapping relationship between various systems and standards, and the realization of data conversion through ETL,

the standardization of these two types of data can be solved. Then, modeling is carried out according to topics such as workload, department business, and nursing evaluation, and the necessary dimensions are divided into these topics. The entire design consists of several sets of fact tables and dimension tables. The fact table stores the calculated data with less data redundancy and is directly provided for data mining. However, the generation of the fact table is based on a large number of calculations. The dimension table records different aspects in the fact table, each dimension table represents an observation angle, and the attributes in the dimension table represent a finer granularity, which can usually support further drilling of the data. At the same time, in order to improve the efficiency of data processing and reduce the impact on the business system, a corresponding temporary table (Stage) is established according to the structure of the business system table. The structure of the Stage table is as consistent as possible with the business system. Figure 2 shows the design of this data warehouse modeling.

### 3.4. Design of Data Mining Model.

The fundamental purpose of the establishment of a data warehouse is to support data mining and fulfill the needs of the hospital. Therefore, the design of this data mining model follows the following principles. First, we must build a new incentive performance pay model and replace income with the workload. After the model is established, historical data should be used to simulate and calculate. The total amount of incentive performance pay is equal to the historical total, and the incentive performance pay of each month is as close as possible to the previous ones. Second, it is necessary to solve the problem of division of departments and use a more scientific method to group and sort the various clinical departments of the hospital so that the hospital can determine the performance adjustment and incentive performance pay plan.

### 3.4.1. Design of Workload Performance Model.

In the past, the performance model used by hospitals was mainly based on the balance of income and expenditure. In simple terms, it is (income-expenditure) × coefficient. The determination of the coefficient is determined by the hospital in accordance with internal discussions, experience reference, and peer reference. To change such a method, the first step is to establish how to assess workload performance. From the analysis of the doctor's daily work, the main items include surgery, treatment, billing, handling of medical records, rounds, and outpatient services. The former data can be collected by the HIS system. The HIS system includes billing and execution projects of the whole hospital. The latter can be obtained from the medical record management system. In accordance with the requirements of the higher-level health system, the medical record system needs to fill in the health statistics on a monthly basis, and this part of the data can be used. Starting from the hospital management goals, the incentive performance pay for this plan was formulated. 70% comes from the medical projects that doctors participate in,
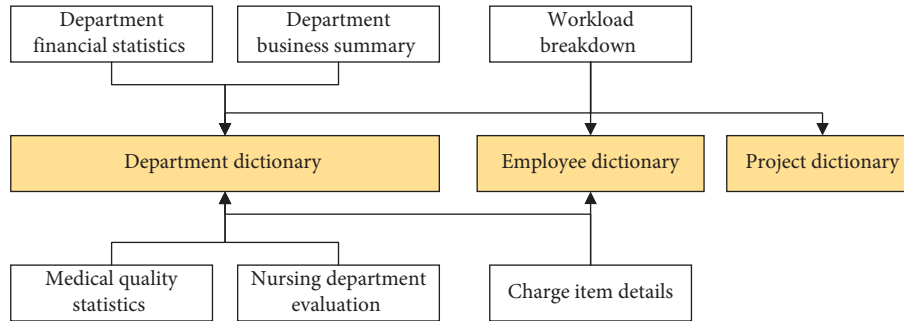
FIGURE 2: Model of the data warehouse.

and 30% comes from the number of patients that doctors participate in treatment.

The workload of doctors is jointly assessed from two aspects. It is easy to understand about outpatients and inpatients, but the medical items that are billed and executed are quite special. It is allowed to use billing to calculate the performance of doctors. In terms of the total workload of doctors and nurses, the income of the project is not used as the accounting standard, and the RBRVS points in the United States are used as the basis for calculating each treatment item. Each project in the HIS system is given a new point. Using the points and the number of projects executed by each department, you can calculate the number of project points for each department and each month. Work performance points are as follows. According to the hospital's proportional requirements, the historical monthly incentive performance pay is divided into 70% and 30%, and 70% of the incentive performance pay for each department and each month is used as the basic incentive performance pay for the performance part of the project, and the remaining 30% is used as the patient basic incentive performance pay in terms of attendance. In order to better follow the change of department incentive performance pay within the controllable range, the program adopts a one-variable linear regression equation as the modeling of the incentive performance pay calculation of each department. Taking the performance pay measurement of the project part as an example, using a linear regression equation to establish the relationship between the workload and performance pay variables, and using data samples for analysis and prediction, the formula is as follows:

$$y = \beta_0 + \beta_1 * x. \tag{1}$$

### 3.4.2. Design of Division Model.
In order to ensure the smooth progress of the performance reform, 3% of the hospital's income will be used as a supplement to the expected incentive performance pay. However, this part of the incentive performance pay is not evenly distributed to each department but is based on the value of the discipline, labor intensity, and other aspects. It needs to be targeted at key disciplines, high-tech departments, and labor-intensive departments that are not willing to go. Or it is similar to the

low-profit departments of Chinese medicine massage to give a certain degree of tilt and care.

To solve this problem, we should first classify the departments based on comprehensive consideration of all aspects of the departments and then give the departments a certain degree of support based on the classification. However, any subjective division is unreasonable. In most cases, it is only possible to classify departments through a certain index, but it is difficult to make a direct judgment. Therefore, it is necessary to provide a quantifiable classification for the hospital through the method of data mining and then establish the classification of the departments with the help of clinicians and nurses' subjective understanding of each department. Cluster analysis is different from the classification algorithm. There is no label to mark the grouping objects in advance, so the cluster analysis is especially suitable for grouping departments. In this scheme, the $K$-means algorithm in cluster analysis is used to first classify the nursing units, which are divided into 6 clusters in total, namely, 6 major types of nursing units, and then the professionals judge the class and the level of the class. In order to evaluate nursing units more objectively, departments must be evaluated from both subjective and objective perspectives. The so-called objective refers to the clinical data recorded by various departments in the HIS, medical records, and other systems. This type of data does not depend on someone's will; the so-called subjective perspective refers to the use of expert scoring methods to score nursing units.

### 3.5. Data Mining Design Based on Weka.
This project requires the implementation of data mining functions in the performance management system. After understanding the software in related fields, we decided to use weka for internal integration. Through the performance management software, we call the interface of weka and use existing algorithms to meet the above data-mining design requirements.

### 3.5.1. Design Based on One-Variable Linear Regression Equation.
In order to obtain data more efficiently, the data acquisition of the data warehouse is provided by the database stored procedures. The sequence diagram of calling weka to solve the unary linear regression equation in the system is shown in Figure 3.
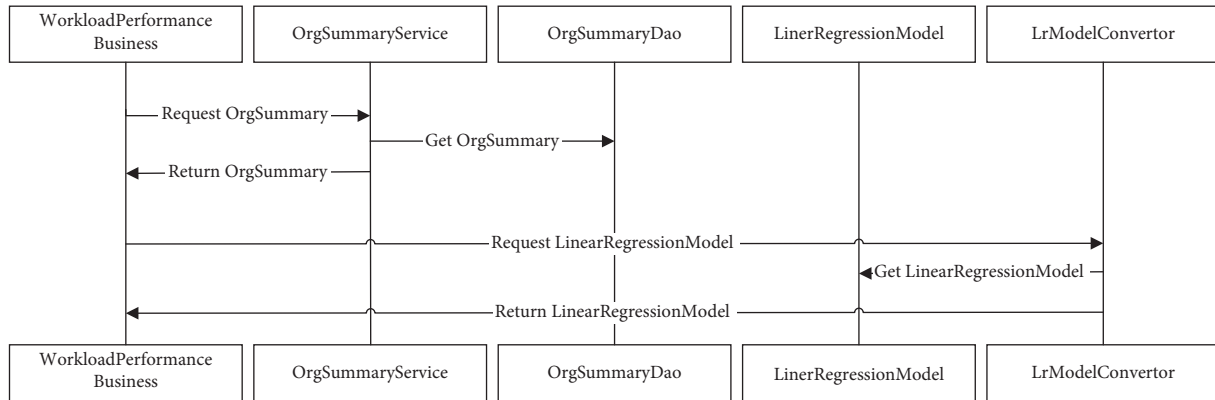
Figure 3: SD of implementation of linear regression.

To integrate unary linear regression in the performance management system, first, it is necessary to obtain a collection of corresponding tables in the data warehouse through the Service layer and convert these collections into a Dao class collection in the form of a List. According to the definition of performance appraisal, at least three types of data are required: a collection of project points related to workload, a collection of medical record data related to the number of patients treated, and a collection of historical incentive performance pay related to incentive performance pay. All the required data can be obtained directly through OrgSummaryService, which is OrgSummaryDao. The Business layer is in charge of business logic, interprets the requirements at the application layer and issues them to the Service layer, or calls the Service layer to obtain the required persistent data. The Business layer of this system obtains the three types of Dao data that they are responsible for by calling the Service layer and then combines and logically transforms them inside the Service to generate the Arff format that weka can recognize. The Business layer finally distributes the Arff format to a tool class such as LrModelConvertor. The tool class converts Arff into a human-recognizable LrModel and returns it to the Business layer.

### 3.5.2. Design Based on K-Means.
The same as the design idea of linear regression, the top-down structure such as Business, Service, and Dao is still adopted, and it depends on downward. The sequence diagram is shown in Figure 4.

NurseClassifyBusiness is used to call the application layer and is responsible for processing business logic and responding to the request of the application layer to call KMeansModel. NurseClassifyBusiness passes the request to two Services: NurseAssessmentService and OrgSummaryService. After obtaining their Dao, respectively, they return the Dao to NurseClassifyBusiness. NurseClassifyBusiness internally processes the conversion logic, generates a string in Arff format that weka's simple K-means can recognize, and then passes the string to KMeansModelConvertor. As an independent conversion class, KMeansModelConvertor is responsible for converting Arff format strings into a KMeansModel structure, in which the parameters required by SimpleKMeans will

be set, and then weka's API will be called, and the result of weka's API will be processed to generate the required KMeansModel.

## 4. Evaluation and Discussion

After designing and implementing the performance management system and implementing the data mining function, the system is further used and evaluated.

### 4.1. Design Ideas of Test Cases.
Regarding the test cases for data mining verification of this program, we mainly start with several concerns of the hospital. First, whether the total incentive performance pay will exceed expectations, the hospital has added 3% to the total budget of the original performance incentive performance pay ratio. A period of historical data was calculated, and the result was found to be generally higher than this expectation; then, this result is unacceptable. On the contrary, if it can be more generally around this number, the conclusion should be reliable. In the same way, the same concerns exist for each clinical department. Therefore, one of the key points of the test is to ensure that the incentive performance pay of the department and the hospital does not exceed expectations in a historical period and to ensure that the workload and the incentive performance pay are proportional. Second, due to the grouping of departments, whether the grouping algorithm is effective, evaluation can be made from two perspectives. The first is to evaluate the results of the cluster analysis itself and whether the calculated results are mathematically valid. The second is an expert review, through which experts in related fields evaluate the evaluation results to determine whether such an algorithm conforms to the actual situation.

### 4.2. Validation of Workload Performance Model.
In this study, the incentive performance pay model transformed by the unary linear regression equation is used to determine the linear relationship between workload and incentive performance pay. This article will design and verify test cases from the following perspectives.
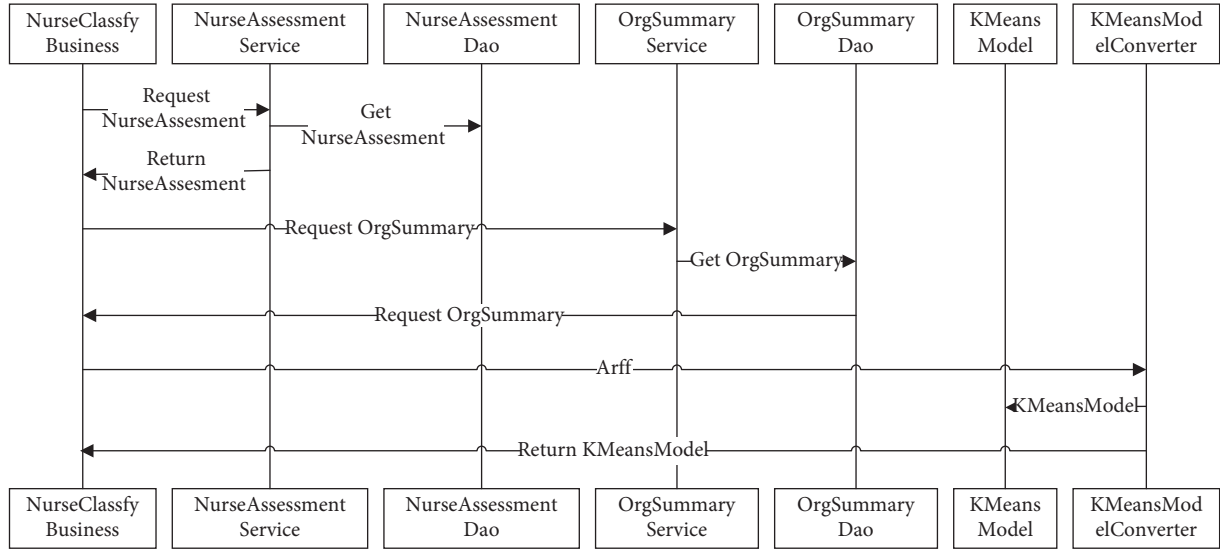
FIGURE 4: SD of implementation of *K*-means.

### 4.2.1. Hospital-Wide Verification.

We perform linear regression analysis on the workload and incentive performance pay of the whole hospital and analyze whether there is a linear relationship between the two from a macroperspective. If the linear relationship between the two can be established, it can be considered that from a global perspective, the new incentive performance pay model can adapt to the actual situation of the hospital. At the same time, because the fluctuation of the whole hospital incentive performance pay is smaller than that of the department, the requirement for the correlation coefficient $r$ is slightly higher, and it is believed that it should reach above 0.85, which means that the expected value of the test case should reach 0.85. The formula of the hospital-wide incentive performance pay model is Hospital-wide incentive performance pay = $A \times$ the workload of the whole hospital + $B$. After calculation, the correlation coefficient of the whole hospital group of physicians is shown in Table 1.

Among them, $r$ is the correlation coefficient. It can be seen that $r$ reaches 0.88. It can be considered that the workload of the whole hospital is highly correlated with the incentive performance pay, which can better reflect the linear relationship between the workload of the whole hospital and the incentive performance pay.

### 4.2.2. Department Verification.

By testing a single department, it is determined whether its linear regression equation can be established, that is, whether the linear relationship between workload and incentive performance pay is established. One of the intentions of the performance model of this project is to solve the relationship between workload and incentive performance pay, and incentive performance pay must be based on history. In order to achieve the same total amount, using the averaging method to solve the problem can also achieve the same effect; that is, the total amount in the same period in history is the same. However, in order to further make the historical incentive

TABLE 1: Correlation coefficient for doctor group.

| Item | $A$ | $B$ | $r$ |
|---|---|---|---|
| Hospital | 4.87 | 490315 | 0.87 |

performance pay close to the actual value, this project chose linear regression equation modeling. Because historical incentive performance pay is affected by costs, the National Chiao Tung University had to relax the requirements on the value of $r$ during the calculation. It is believed that the linear equation is still valid even when $r$ is between 0.5 and 1, and the expected value of the test case should reach 0.5 or more. The formula of the incentive performance pay model is Department incentive performance pay = $A \times$ department workload + $B$. Table 2 is the linear regression solution for some physician departments in the hospital.

### 4.3. Validation of Department Grouping Model.

The model verification of the division of departments should be designed from two aspects: first, the significance of the clustering algorithm itself, that is, how to evaluate whether the conclusions of the clustering algorithm are mathematically significant; the second is the evaluation of the results and the analysis by the expert team, to ensure that the grouping results have practical business significance.

### 4.3.1. Clustering Result Evaluation.

We use the K-means model in the weka software to classify nursing units. The software provides us with a piece of data that can assist us in judging the effect of clustering. The size of the value represents the spatial distance between instances in the same cluster. The smaller the value, the smaller the distance, the smaller the distance, the better the effect. By adjusting the number of seeds, different results can usually be obtained, but based on the nature of K-means, the results are often only locally optimal solutions. In this scheme, first, we use the default seed 10 to perform clustering calculations on the

TABLE 2: Correlation coefficient for organizations.

| Item | A | B | r |
| --- | --- | --- | --- |
| Neurosurgery | 8.01 | 2993 | 0.83 |
| Orthopedics | 3.45 | 13972 | 0.41 |
| Urology | 8.56 | 680 | 0.68 |
| General surgery | 6.93 | 11201 | 0.82 |
| Cardiac surgery | 3.57 | 11767 | 0.34 |
| Thoracic surgery | 7.60 | 2545 | 0.80 |
| Endocrinology | 3.76 | 12805 | 0.89 |
| Hematology | 2.21 | 21098 | 0.48 |

Among them, $A$ is the coefficient of workload, $B$ is the constant of the department, and $r$ is the correlation coefficient. It can be found that there is still a part of the correlation coefficient that does not meet expectations.

nursing department, and the result of this indicator is 83.653. Subsequently, the number of seeds was adjusted from 1 to 200. The best result is that when the number of seeds is 139, the index is 82.582. Therefore, it can be considered that the result of this clustering is ideal, and the result of this time is adopted.

*4.3.2. Expert Evaluation.* The expert team organized by the hospital includes the head nurses of various departments, the director of the nursing department, and the rehired experts. The expert team conducted a human analysis of the calculation results and found that departments with high work intensity and high technical difficulty can be effectively classified into one category, and two departments with good hardware conditions and working environment and low work pressure can also be classified into one category. The expert team unanimously recognized the calculation results and believed that the calculation results are effective and can effectively reflect the level of the department.

*4.4. Analysis and Evaluation of Test Results.* Through the above verification results, it can be found that the results obtained by data mining are basically consistent with the expected results, and there are certain errors, but the final cause of the deviation can also be found. After analysis, the error is considered to be within the acceptable range. The performance reform project finally ended with a smooth acceptance.

## 5. Conclusion

The contributions of this paper are as follows:

(1) After the project of the hospital performance management system, we understand the characteristics of the medical industry and are familiar with the hospital business process, and combined with the focus of the hospital staff, we established a new workload-based incentive performance pay assessment system. At the same time, combined with data mining technology, objective and credible data provide a theoretical basis for the hospital to help the hospital further implement performance reforms.

(2) We researched the related algorithms of data mining, flexibly used the univariate linear regression equation and cluster analysis in the project, helped the hospital establish an incentive performance pay model, and completed the department classification work to provide a basis for the formulation of performance policies.

(3) We are familiar with the theory and related technologies of data warehouse, use the idea of data warehouse for modeling and dimension division, use SQL Server stored procedures to create scripts, implement ETL by deploying database tasks, and effectively assist in data mining.

(4) In the hospital performance management system, in addition to the data mining function, there are many business applications, such as permissions, incentive performance pay calculation, and report display, which are not detailed in this article, but only we explain the software functions of data mining. As an important technical method in the consulting industry, data mining can help hospitals analyze and solve problems well. Therefore, in this article, we will focus on data mining.

## Data Availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

## Additional Points

We provide further familiarity and understanding of the weka tool, are able to realize the corresponding data mining algorithm through the learning of weka's API, and also lay the foundation for the future use of weka to achieve more functions.

## Disclosure

Huiqun Lu and RuiLing Wang are co-first authors.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

Huiqun Lu and RuiLing Wang contributed equally to this work.

## References

[1] X. Hou, "Analysis on hotspots of venture investment in the medical and health industry under the new medical reform policy," vol. 27, p. 146, 2009.

[2] W. Arnett, *Computerized Medical and Dental Diagnosis and Treatment Planning and Management System*, 2005.

[3] S. L. Ting, S. K. Kwok, A. H. C. Tsang, and W. B. Lee, "Critical elements and lessons learnt from the implementation of an

RFID-enabled healthcare management system in a medical organization," *Journal of Medical Systems*, vol. 35, no. 4, pp. 657–669, 2009.

[4] J. D. Trachtenberg, A. D. Kambugu, M. Mckellar et al., "The medical management of central nervous system infections in Uganda and the potential impact of an algorithm-based approach to improve outcomes," *International Journal of Infectious Diseases*, vol. 11, no. 6, pp. 524–530, 2007.

[5] V. Martinez, F. Alberto, A. Fornasier, M. Montagner, and F. Nonino, *International Journal of Productivity and Performance Management*, vol. 56, no. 5/6, pp. 417–435, 2007.

[6] J. Pei, "Solving the problem of charging and discharging of electric vehicles based on particle swarm algorithm," in *Processings of the International Conference on Information Systems and Computer Aided Education*, pp. 534–538.

[7] Y. Xing, H. Wei, J. Suo et al., "Construction of occupational safety and health management laws and regulations and system in medical service," *Chinese Journal of Nosocomiology*, vol. 22, 2006.

[8] H. Xu, "Design and research of hospital medical supplies management information system," *Chinese Journal of Medical Instrumentation*, vol. 34, 2009.

[9] D. Otley, "Performance management: a framework for management control systems research," *Management Accounting Research*, vol. 10, no. 4, pp. 363–382, 1999.

[10] S. Sharma, "International journal of productivity and performance management," *International Journal of Productivity and Performance Management*, vol. 63, no. 1, pp. 57–84, 2006.

[11] B. Aldonio, "The design and use of performance management systems: an extended framework for analysis," *Management Accounting Research*, vol. 20, no. 4, pp. 263–282, 2009.

[12] K. Fryer, J. Antony, and S. Ogden, "Performance management in the public sector," *International Journal of Public Sector Management*, vol. 22, no. 6, pp. 478–498, 2013.

[13] S. Parekh, N. Gandhi, J. Hellerstein et al., "Using control theory to achieve service level objectives in performance management," *Real-Time Systems*, vol. 23, no. 1-2, pp. 127–141, 2002.

[14] D. C. S. Swarnalatha, T. Prasanna, and T. S. P. Prasanna, "Performance management and employee engagement," *Indian Journal of Applied Research*, vol. 2, no. 3, pp. 132–136, 2011.

[15] L. Zhong and S. Kan, "Several basic issues of performance management," *Nankai Business Review*, vol. 12, 2002.

[16] K. Zhong, P. Wang, and J. Pei, "Multi objective optimization regarding vehicles and power grids," *Wireless Communications and Mobile Computing*, vol. 23, 2021.

[17] J. Wen, "Analysis on the development trend of hospital performance management," vol. 12, p. 60, 2013.

[18] V. Gueorguieva, J. Accius, C. Apaza et al., "The program assessment rating tool and the government performance and results act: evaluating conflicts and disconnections," *The American Review of Public Administration*, vol. 39, no. 3, pp. 225–245, 2009.

[19] J. Yang, L. Fang, and W. Lin, "Thoughts on several basic issues of performance evaluation," *Journal of Dialectics of Nature*, vol. 23, no. 132, pp. 40–51, 2001.

[20] C. Huang, *Medical Institution Management System*, Shanghai Jiao Tong University Press, Shanghai, China, 2006.

[21] X. Ma, "Establishment of performance management index system of public hospital," *Hosp Admin J Chin PLA*, vol. 19, no. 7, pp. 608-609, 2012.

[22] M. Bagaria and A. Facs, *Key Performance Indicators: Satisfying Quest for Quality*, Journal of the American College of Surgeons, Chicago, Illinois, United States, 2021.

[23] J. Wu, J. Ye, F. Gao et al., "Practice and experience of RBRVS assessment system on physician performance management," *Chinese Hospitals*, vol. 17, no. 2, pp. 49–51, 2013.

[24] J. Li, L. Tian, H. Xiu et al., "Hospital strategic performance evaluation index system," *Chinese Hospital Management*, vol. 32, no. 3, pp. 25–27, 2012.

[25] Y. Han and L. Cao, "Analysis of the effect of the introduction of Taiwanese hospitals' performance management," *Hospital Management Forum*, vol. 27, no. 4, pp. 41-42, 2010.

[26] Z. Wu, "Research and application of web log mining," *Science Mosaic*, vol. 6, pp. 43–45, 2010.

[27] Y. Xia, J. Zhou, and C. Xiang, "Research progress in modern data mining technology," *Acta Agriculture Jiangxi*, vol. 21, no. 4, pp. 82–84, 2009.

[28] J. Pei, "Big data mining in the control of epidemic," *Basic and Clinical Pharmacology and Toxicology*, vol. 428-430, 2020.