UDC 621.396.4(045)
DOI: 10.18372/2306-1472.87.15673

Oleksandr Davydko[1]
Liliia Levchyk[2]
Olexiy Khlimenkov[3]
Olga Averianova[4]

DATA STRUCTURE OF MEDICAL INFORMATION ORGANIZATION AND SELECTING OPTIMAL PATIENT TREATMENT STRATEGIES

National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute»,
37, Peremohy ave., Kyiv, 03056, Ukraine
E-mails: [1]alexander.davydko@gmail.com; [2]angelicaannalilia@gmail.com; [3]xlimenkov4@gmail.com; [4]olgaaveryanova@ukr.net

Abstract

Purpose: *The purpose of this article is to develop a clinical information storage and medical decision support system for daily use by physicians during their work with patients.* ***Methods:*** *The article considered approaches to achieve the research goal through the creation of a backend on Django Rest Framework, as well as the application of the simplex method to perform treatment optimization.* ***Results:*** *A server-side software product was obtained, where a special database was implemented, which takes into account the requirements for the data structure of the organization of medical information, and a module for selecting optimal treatment strategies for patients was created.* ***Discussion:*** *The conducted research allowed us to obtain a prototype of medical information system, which will be useful due to its versatility in the data structure of the organization of medical data and the use of machine learning methods to optimize treatment strategies.*

Keywords: data organization, database structure, treatment optimization, treatment strategies, data export.

1. Introduction

Health care is one of the main areas of life, which requires attention and investment in various ways, including the introduction of information technology to improve and optimize the processes of medical institutions. Automation of medicine, implementation of medical information systems (MIS), and electronic document management in health care institutions allowed to improve the quality of medical services.

With a large number of available software products for implementation in the Ministry of Health on the market, the question of developing a universal structure of the organization of medical data taking into account all the details of the treatment process remains open. In addition, systems that not only provide basic functions for working with medical data (such as storing, updating, and writing to the database) but also provides the ability to independently analyze and solve problems by creating mathematical models based on available data using machine-learning algorithms are especially popular. This allows the human error factor to be eliminated as much as possible when making decisions on patient treatment strategies and to predict and respond to possible health threats.

2. Analysis of the latest research and publications

Modern developments in the field of information technology (IT) have enabled the development of new ways of organizing and providing medical services. Saving and exchange of medical information between different health care institutions, online teleconsultations, remote monitoring of the patient's condition, statistical analysis of all information in the database to identify new trends and needs of the population - all this and much more has become possible due to the development of modern technology. To ensure that the MIS meets all the necessary functional requirements, it is necessary first of all to pay attention to the technology of organizing and storing data. At the same time, it is necessary to take into account the requirements for adaptive data structures, business logic, and technologies of designing MIS user interface, which accordingly dictates their own rules for methods and formats of patient data presentation.

A centralized architecture was developed for the MIS, which can be used both for one MIS and for the whole range of MISs. The main thing in this approach is the functioning of a single data repository or the central database (CDB) on the same server capacity. Since the MIS includes different types of data, it is logical to classify the data, which was considered in the work of Russian scientists [1]. The following classes of data are distinguished:

- Structured data of simple types, for example, dates, numbers, and rows. IT specialists are used to operating such data in relational databases.
- Partially structured data are documents or objects with the structure of which the MIS work is organized.
- Non-structured data are often binary objects in various formats, such as graphical images (photos of ultrasound investigation, computer, and magnetic resonance imaging, electrocardiography, etc.), scanned documents, and documents of various editors containing information necessary for the work of MIS.

In addition, all of the above-listed data types have a different purpose. Some of them contain medical records and data about the patient's disease history, diagnosis, duration of the disease, medications prescribed and treatment measures taken, analyses, and patient's condition; Other data are intended for general information about the patient, such as age, gender, place of work, addresses, contacts, social and family status. They are not directly related to the treatment process but are important for the physician's review and understanding of the patient's holistic picture.

Regarding data presentation, relational variants are most often used as database management systems (DBMS), non-relational variants are less common, and some MISs have taken the documentary approach as the basis. Anyway, none of the above options in a pure form meet the conditions for presentation and preservation of patient data, therefore in the end the developers decided to use a mixed type of information representation [2-3], for example, object-relational or relational-object, depending on the MIS technological platform. The specialists of Interin, who developed the "Interin PROMIS" MIS [4], also concluded the need for a combination of data representation [3].

Nowadays, at the initiative of the Ministry of Health of Ukraine, *eHealth*, an electronic system of health care, continues to be implemented in medical institutions, which includes the CDB and MIS of medical institutions. Most health care institutions have fully or partially switched to the electronic system of patient care, and there are nearly 41 MISs connected to eHealth, using all of its functionality, or some of its parts. It should also be noted that the active development and improvement of the eHealth system is underway to increase its effectiveness and compliance with the requirements of business processes. The system also plans to expand its functionality by adding new modules of reimbursement, sensitive data, treatment plan, and electronic medical treatment, as well as an electronic patient's office.

3. Problem statement

The purpose of this study is to develop a web additive structure of medical information organization data, implementation of the data export function in the format, which is convenient for further analysis of such data using machine learning methods, and the function of selecting the optimal treatment strategies for the patients and predetermining their condition after the treatment strategy is reversed. The implementation of this functionality in its version of the system will not only provide a new counterpart in the Ukrainian market of MIS but also make it possible to fully use modern methods of machine learning in medicine.

As for the implementation of the function of selecting optimal treatment strategies for patients, it is conditioned by the fact that physicians, who independently decide on the treatment strategy, for their specialty do not have the full baggage of knowledge of all types of drugs and their appropriate use. This duty lies on the clinical pharmacist, but this profession in Ukraine is not developed, and it is quite rare to find such a practitioner in medical institutions. Therefore, automation of this process is also an urgent task.

4. Problem solution

For the development of the hardware and software part of this system (backend), we used the Python programming language, as well as the Django Rest Framework for the development of web services. PostgreSQL was used as a DBMS. The implementation of the structure of medical information organization was carried out according to the plan described in table.

Table

Project development and implementation plan

Stage	Contents of the work	Result
1	Designing database structure and data dependencies	Graphical implementation of the database structure in the form of tables and fields with the appropriate links
2	Creating the project in the software environment, setting up the basic functions of the project, and creating models according to the database	Project files in *Python* programming language using *Django Rest Framework*, which contain basic settings and database table models
3	Creating an Application Programming Interface (API)	The ability to run the server and display the structure and model data on the backend page using API
4	Filling models with test data and connecting to the *PostgreSQL* database	Synchronization with DBMS, recording and reading data from the database
5	Implementation of user requests on the local server via API and database	Displaying the required information when entering the corresponding URL in the browser
6	Implementation of exporting patient data to an *Excel* file	Excel file with data about the patient's disease history
7	Implementation of the mechanism of analysis and predetermination of treatment results using machine learning in scikit-learn.	The ability of the system to build models by the method of group method of data handling (GMDH) and to predict the results of patient treatment.

4.1. Data structure of medical information organization

The developed structure of medical information organization allows to structure the data of both the patient and medical staff, as well as to store detailed information about every stage of patient treatment in the hospital. Figure 1 shows the structure of the database where medical information is stored.

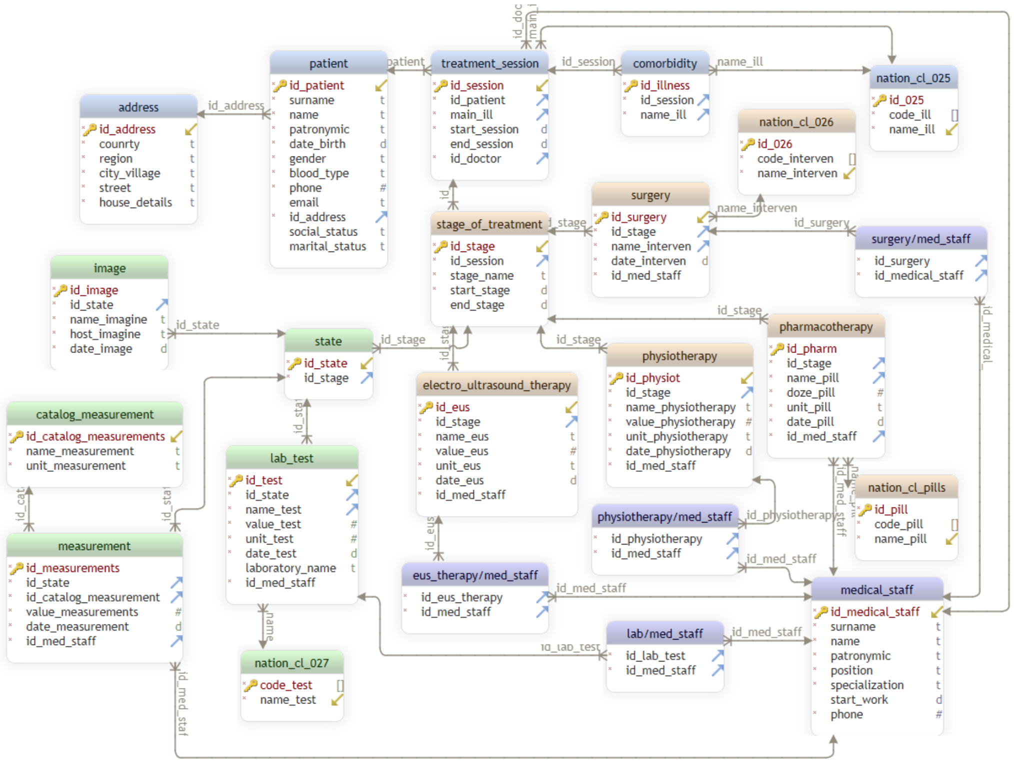

Fig. 1. Database structure

The system allows storing general information about the patient (e.g., name, surname, date of birth, age, blood type, contact data, social status) as well as data about his/her condition and treatment stages organized in the input structures.

Thus, for each patient, information about the treatment session is stored (and there can be several sessions), which, in turn, includes data about the patient's related diseases and treatment stages at the time of the treatment session.

The treatment phase also includes such inputs as the patient's condition at the given stage of treatment, including various treatment methods:

- surgery;
- pharmacotherapy;
- physical therapy;
- electro-ultrasound therapy.

In addition, the patient's condition at each stage of the treatment session is determined by the following components: image scans, laboratory tests, measurements of the patient's condition (e.g., pulse, arterial blood pressure, temperature, weight, height). Special attention should be paid to the inclusion of national classifiers into the database, such as:

- National Classifier 025:2019 "Classifier of Diseases and Controversial Health Problems" [5].
- National Classifier 026:2019 "Medical Intervention Classifier" [6].
- National Classifier 027:2019 "Classifier of Laboratory Tests and Indicators" [7].
- National Classifier 024:2019 "Classifier of Medical Products" [8].

It should also be noted that the system allows to save information not only about patients, but also about medical staff, and any actions of medical staff are recorded in the corresponding table of the stage of treatment or condition of the patient. These links can be seen in Fig. 1.

4.2. Calculation of optimal patient treatment strategies

To implement the selection of optimal treatment strategies using machine learning and testing the performance of this module of the system, we used a database of children with cardiac problems who were treated in the cardiology department for diagnoses of aortic stenosis (reduced density of the aortic arch to the aorta, which complicates blood flow from the left ventricle to the aorta) and aortic stenosis: Aortic stenosis (decreased density of the aortic orifice, which impedes blood flow from the left ventricle to the aorta) and aortic valve anomaly and/or obstruction of the left ventricle.

The database was provided by the Amosov National Institute of Cardiovascular Surgery. This database contains general parameters of the patients, results of their analysis and diagnostics before and after treatment, and information on treatment methods and doses for each patient. For the study the following parameters were taken from the database:

- age of the patient (x_1);
- weight (x_2);
- end-diastolic pressure before treatment (x_3);
- aortic pressure gradient before treatment (x_4);
- augmentation index before treatment (x_5);
- emission fraction before treatment (x_6);
- dose of the drug A (u_1);
- dose of the drug B (u_2);
- end-diastolic pressure after treatment (y_1);
- pressure gradient on the aorta after treatment (y_2);
- augmentation index after treatment (y_3);
- emission fraction after treatment (y_4).

The names of drugs A and B were hidden due to the risk of their use by third-party users for self-medication.

The task of choosing the optimal treatment strategy is to find the dosage of drugs that will give the optimal patient's condition during treatment using the optimization method [9]. The technique for solving such problems was proposed in the work [10].

The information system created in this work allows using the described technology in automated mode. The first step is to build statistical models for the criterion that determines the quality of treatment and patient end states (y_1, …, y_4), where the inputs will be treatment parameters (x_1, …, x_6) and medications (u_1 i u_2). In this particular task these models will look like:

$$\begin{cases} y_1 = a_1^1 x_1 + ... + a_6^1 x_6 + a_7^1 u_1 + a_8^1 u_2 \\ y_2 = a_1^2 x_1 + ... + a_6^2 x_6 + a_7^2 u_1 + a_8^2 u_2 \\ y_3 = a_1^3 x_1 + ... + a_6^3 x_6 + a_7^3 u_1 + a_8^3 u_2 \\ y_4 = a_1^4 x_1 + ... + a_6^4 x_6 + a_7^4 u_1 + a_8^4 u_2 \end{cases} \quad (1)$$

where: a_1, …, a_8 – coefficients at the input variables in the model.

Models y_1, …, y_4 are obtained by means of algorithms of the group method of data handling (GMDH). For the purpose of using the most effective apparatus of mathematical programming we will search for models in the class of linear by u_1 and u_2

Below is the demonstrated application of nonlinear by parameters x and linear by u model of the criterion, the coefficient of determination (R2) of which is 0.877:

$$y_1 = u_1\left(\frac{0.284x_1}{x_6} - \frac{1.043x_2}{x_6} - \frac{0.013x_2}{x_4}\right) + \\ u_2\left(0.872 - \frac{0.223x_2}{x_4} + \frac{1.485x_5}{x_3} - \frac{0.167x_2}{x_1}\right) + \\ + \frac{3.948x_1}{x_2} - 10.498 \qquad (2)$$

The models (1) are used to form the optimization problem:

$$\begin{cases} a_1^1 x_1 + ... + a_6^1 x_6 + a_7^1 u_1 + a_8^1 u_2 \to \min \\ c_{\min}^2 \le a_1^2 x_1 + ... + a_6^2 x_6 + a_7^2 u_1 + a_8^2 u_2 \le c_{\max}^2 \\ c_{\min}^3 \le a_1^3 x_1 + ... + a_6^3 x_6 + a_7^3 u_1 + a_8^3 u_2 \le c_{\max}^3 \\ c_{\min}^4 \le a_1^4 x_1 + ... + a_6^4 x_6 + a_7^4 u_1 + a_8^4 u_2 \le c_{\max}^4 \end{cases}. \qquad (3)$$

It is assumed that parameters x_1, …, x_6 are specified by the order (since the patient was hospitalized with them) when selecting the optimal treatment strategy. Therefore, problem (2) can be rewritten as follows:

$$\begin{cases} a_7^1 u_1 + a_8^1 u_2 + const \to \min \\ c_{\min}^2 \le a_7^2 u_1 + a_8^2 u_2 + const \le c_{\max}^2 \\ c_{\min}^3 \le a_7^3 u_1 + a_8^3 u_2 + const \le c_{\max}^3 \\ c_{\min}^4 \le a_7^4 u_1 + a_8^4 u_2 + const \le c_{\max}^4 \end{cases}, \qquad (4)$$

where: *const* – value, which was obtained after inserting the values of the parameters x_1, …, x_6. The implementation of this approach was done in the following way: in *Django* framework models, a separate class was created to describe the parameters that will come to the input of the model with the specified field types of these parameters by adding the external key of the Patient table, to uniquely identify the person for whom the optimal treatment strategy will be calculated, and search for the values of these parameters in the database. After the doctor sent a request for optimization by specifying the parameters and patient id (identifier), the database is searched by the *Patient* table where the id key is stored and by the *Measurement* and *LaboratoryTest* tables, which contain the names of the patient's condition parameters.

The next step is to solve the formulated optimization problem. This function analyzes the input values and searches for possible treatment options by using the methods of the library *sklearn* (*scikit-learn*). Among the potential treatment, strategies are selected those that have the best indicators of the criterion (*strategy_score*) and do not go beyond the limits on the end states of the patient. The result of the calculations was sent back to the API, where the doctor can find information about the treatment strategies proposed by the system. To provide this data in a readable format, we also wrote a "serializer" for model parameters, which, with the help of DRF tools and class viewsets, provides the user with information about the calculation of optimal treatment strategies and patient's condition parameters in json format.

5. Conclusions

As a result of this research, the structure of medical information organization data and selection of optimal treatment strategies for patients was developed. For this purpose, we created a server-side software product and used the technologies listed in Table 1.

To obtain the data structure, we created the database scheme shown in Fig. 1, as well as used national classifiers [5-8], which are an integral part of any medical information system in Ukraine.

To implement the module for selecting optimal treatment strategies for patients, we used the approach that described in more detail in the work [11] by using the simplex method as an optimization algorithm. The Django framework used for software implementation.

References

[1] D. V. Belishev, E. V. Kochurov. "Perspective Methods of Data Management in Medical Information Systems". 2016, pp. 79-97.

[2] A. G. Belishev, J.I. Guliev, V.Y. Morozov. "Building Medical Systems with Object Technologies," Program Systems: Theoretical Foundations and Supplements, ed. by A. K. Ailamazyan, Nauka. Fizmatlit, M., 1999, 169 p.

[3] V. L. Malyh, S. P. Pimenov, M. I. Khatkevich. "The Object-Relational Approach to the Creation of Large Information Systems," Program Systems: Theoretical Foundations and Supplements, ed. by A. K. Ailamazyan, Nauka. Fizmatlit, M., 1999, 177 p.

[4] Y. I. Guliev, S.I. Komarov, V.L. Malyh, G.S. Osipov, S.P. Pimenov, M.I. Khatkevich "Integrated disposable information system Treatment Facility (INTERN)", Software Products and Systems, 1997, №3.

[5] National Classifier 025:2019 "The Classifier of diseases and controversial health problems" Available at: https://bit.ly/3tRdPD0

[6] National Classifier 026:2019 "Medical Interventions Classifier" Available at: https://bit.ly/3tRdPD0

[7] National Classifier 027:2019 "Classifier of laboratory tests and indicators" Available at: https://bit.ly/3eP5vPB

[8] National Classifier 024:2019 "Medicinal Products Classifier" Available at: https://bit.ly/2SVJ7Mb

[9] Yang X.-S. Introduction to mathematical optimization / X.-S. Yang. — Cambridge International Science Publishing, 2008. — 150 p.

[10] Nastenko I, Pavlov V, Nosovets O, Zelensky K, Davidko O, Pavlov O. Solving the individual control strategy tasks using the optimal complexity models built on the class of similar objects. In: Advances in Intelligent Systems and Computing. Springer; 2020, pp. 535–546.

О.Б. Давидько[1], Л.О. Левчик[2], О.Л. Хліменков[3], О.А. Аверьянова[4]

Структура даних організації медичної інформації та вибір оптимальних стратегій лікування пацієнтів

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", пр. Перемоги, 37, Київ, Україна, 03056
E-mails: [1]alexander.davydko@gmail.com; [2]angelicaannalilia@gmail.com; [3]xlimenkov4@gmail.com, [4]olgaaveryanova@ukr.net

Мета: Метою даної статті є розробка системи зберігання клінічної інформації та підтримки прийняття медичних рішень для щоденного використання лікарями під час їхньої роботи з пацієнтами. **Методи:** У статті були розглянуті підходи для досягнення мети дослідження через створення бекенду на *Django Rest Framework*, а також застосуванням симплекс-методу для виконання оптимізації лікування. **Результати:** Було одержано програмний продукт серверної частини, де реалізована спеціальна база даних, яка враховує вимоги до структури даних організації медичної інформації, а також створено модуль вибору оптимальних стратегій лікування пацієнтів. **Обговорення:** Виконане дослідження дозволило отримати прототип медичної інформаційної системи, яка стане у нагоді завдяки своїй універсальності щодо структури даних організації медичних даних і використанням методів машинного навчання для оптимізації стратегії лікування.

Ключові слова: організація даних, структура БД, оптимізація лікування, стратегії лікування, експорт даних.

А.Б. Давидько[1], Л.О Левчик[2], А.Л. Хлименков[3], О.А. Аверьянова[4]

Структура данных организации медицинской информации и выбор оптимальных стратегий лечения пациентов

Национальный технический университет Украины "Киевский политехнический институт имени Игоря Сикорского", пр. Победы, 37, Киев, Украина, 03056
E-mails: [1]alexander.davydko@gmail.com; [2]angelicaannalilia@gmail.com; [3]xlimenkov4@gmail.com, [4]olgaaveryanova@ukr.net

Цель: Целью данной статьи является разработка системы хранения клинической информации и поддержки принятия медицинских решений для ежедневного использования врачами во время их работы с пациентами. **Методы:** В статье были рассмотрены подходы для достижения цели исследования через создание бэкенд на *Django Rest Framework*, а также применением симплекс-метода для выполнения оптимизации лечения. **Результаты:** Было получено программный продукт серверной части, где реализована специальная база данных, которая учитывает требования к структуре данных организации медицинской информации, а также создан модуль выбора оптимальных стратегий лечения пациентов. **Обсуждение:** Проведенное исследование позволило получить прототип медицинской информационной системы, которая пригодится благодаря своей универсальности по структуре данных организации медицинских данных и использованием методов машинного обучения для оптимизации стратегии лечения.

Ключевые слова: организация данных, структура БД, оптимизация лечения, стратегии лечения, экспорт данных.

Oleksandr Davydko. Postgraduate Student.
Department of Biomedical Cybernetics, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute».
Education: National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» (2019).
Research area: information technologies in medicine, computer science, data science, deep learning.
Publications: 7.
E-mail: alexander.davydko@gmail.com

Liliia Levchyk. Student.
Department of Biomedical Cybernetics, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute».
Research area: information technologies in medicine, computer science, data science, deep learning.
Publications: 0.
E-mail: angelicaannalilia@gmail.com

Olexiy Khlimenkov. Student.
Department of Biomedical Cybernetics, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute».
Research area: information technologies in medicine, computer science, data science, deep learning.
Publications: 0.
E-mail: xlimenkov4@gmail.com

Olga Averianova. Senior Lecturer
Department of Biomedical Cybernetics, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute».
Research area: information technologies in medicine, computer science, data science, deep learning, system analysis, information system design, IT management
Publications: 5.
E-mail: olgaaveryanova@ukr.net