АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР
ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ
INFORMATION AND COMMUNICATION TECHNOLOGIES

DOI 10.51885/1561-4212_2023_3_107
МРНТИ 20.19.21

I.M. Uvaliyeva[1], S.A. Belginova[2], Zh.T. Rakhmetullina[1]
[1]D. Serikbayev East Kazakhstan Technical University, Ust-Kamenogorsk, Kazakhstan
*E-mail: iuvalieva@ektu.kz**
E-mail: zhrakhmetullina@ektu.kz
[2]University «Turan», Almaty, Kazakhstan
E-mail: sbelginova@gmail.com

SCIENTIFIC AND THEORETICAL QUESTIONS OF CLINICAL DECISION SUPPORT TECHNOLOGY BASED ON DIFFERENTIAL DIAGNOSIS OF CLINICAL AND HEMATOLOGICAL SYNDROMES

КЛИНИКАЛЫҚ ЖӘНЕ ГЕМАТОЛОГИЯЛЫҚ СИНДРОМДАРДЫҢ ДИФФЕРЕНЦИАЛДЫ ДИАГНОСТИКАСЫНА НЕГІЗДЕЛГЕН КЛИНИКАЛЫҚ ШЕШІМДЕРДІ ҚОЛДАУ ТЕХНОЛОГИЯСЫНЫҢ ҒЫЛЫМИ-ТЕОРИЯЛЫҚ МӘСЕЛЕЛЕРІ

НАУЧНО-ТЕОРЕТИЧЕСКИЕ ВОПРОСЫ ТЕХНОЛОГИИ ПОДДЕРЖКИ ПРИНЯТИЯ КЛИНИЧЕСКИХ РЕШЕНИЙ НА ОСНОВЕ ДИФФЕРЕНЦИАЛЬНОГО ДИАГНОСТИРОВАНИЯ КЛИНИКО-ГЕМАТОЛОГИЧЕСКИХ СИНДРОМОВ

Abstract. *This study is aimed at developing a software package for diagnosing clinical and hematological syndromes for an electronic passport, which makes it possible to automate the process of clinical decision support based on differential diagnosis algorithms and medical data mining models. The purpose of the study is to develop a software package for an electronic health passport based on the information technology of differential diagnosis and computational-analytical models of the morphological classification of clinical and hematological syndromes. The scientific foundations of the information technology for monitoring clinical and hematological syndromes integrate mathematical models for evaluating, pre-processing and ensemble methods of machine learning of biochemical blood test indicators, as well as methods that improve the efficiency of business processes for remote monitoring of hemoglobin synthesis disorders as a result of iron deficiency developing against the background of various pathological (physiological) conditions and manifested by signs of anemia and sideropenia.*

Keywords: *clinical decision support technology, differential diagnosis, clinical and hematological syndromes, scientific and theoretical base, medical monitoring, information technology, electronic health passport.*

Аңдатпа. *Бұл зерттеу дифференциалды диагностика алгоритмдері мен медициналық деректерді зияткерлік талдау модельдері негізінде клиникалық шешімдерді қабылдауды қолдау процесін автоматтандыруға мүмкіндік беретін электрондық паспортқа арналған клиникалық-гематологиялық синдромдарды диагностикалаудың бағдарламалық кешенін әзірлеуге бағытталған. Зерттеудің мақсаты-дифференциалды диагностиканың ақпараттық технологиясына және клиникалық-гематологиялық синдромдардың морфологиялық классификациясының есептеу-аналитикалық модельдеріне негізделген электрондық денсаулық паспортының бағдарламалық кешенін әзірлеу. Клиникалық-гематологиялық синдромдар мониторингінің ақпараттық технологиясының ғылыми негіздері қанның биохимиялық талдауының көрсеткіштерін бағалаудың, алдын ала өңдеудің және машиналық оқытудың ансамбльдік әдістерінің математикалық модельдерін, сондай-ақ әртүрлі патологиялық (физиологиялық) жағдайлар аясында дамитын және анемия белгілерімен көрінетін темір тапшылығы нәтижесінде гемоглобин синтезінің бұзылуын қашықтықтан бақылаудың бизнес-процестерінің тиімділігін арттыруға мүмкіндік беретін әдістерді біріктіреді және сидеропени.*

Түйін сөздер: *клиникалық шешімдерді қолдау технологиясы, дифференциалды диагностика,*

клиникалық және гематологиялық синдромдар, ғылыми-теориялық база, медициналық мониторинг, ақпараттық технологиялар, электрондық денсаулық паспорты.

Аннотация. *Данное исследование направлено на разработку программного комплекса диагностики клинико-гематологических синдромов для электронного паспорта, позволяющий автоматизировать процесс поддержки принятия клинических решений на основе алгоритмов дифференциальной диагностики и моделей интеллектуального анализа медицинских данных. Цель исследования – разработать программный комплекс электронного паспорта здоровья, основанный на информационной технологии дифференциальной диагностики и вычислительно-аналитических моделях морфологической классификации клинико-гематологических синдромов. Научные основы информационной технологии мониторинга клинико-гематологических синдромов интегрируют математические модели оценки, пре-процессинга и ансамблевых методов машинного обучения показателей биохимического анализа крови, а также методы, позволяющие повысить эффективность бизнес-процессов дистанционного мониторинга нарушения синтеза гемоглобина в результате дефицита железа, развивающегося на фоне различных патологических (физиологических) состояний и проявляющийся признаками анемии и сидеропени.*

Ключевые слова: *технология поддержки принятия клинических решений, дифференциальное диагностирование, клинико-гематологические синдромы, научно-теоретическая база, медицинский мониторинг, информационная технология, электронный паспорт здоровья.*

Introduction. Considering all the recent unfavorable events in the world, including the global pandemic, quarantine restrictions, and economic shocks, the development of the healthcare sector is one of the top priorities of any country.

In this regard, Kazakhstan, like any other country, needs to improve the efficiency of the health sector and the availability of medical care for the entire population, which in turn can be achieved only through digitalization, namely the integration of the main activities of the health sector with information systems, the use of mobile digital applications, the introduction of electronic health passports and the transition to "paperless" hospitals.

For this task, the Ministry of Health, together with experts and representatives of the business community, has developed new approaches to the digitalization of healthcare.

The central element of e-health should be an electronic health passport of every citizen of our country. It is a summary of human health information, including demographic data; information about past and/or existing diseases; biometric data; allergological and immune status.

As part of the digitalization of Kazakhstan's healthcare since 2020, almost all (more precisely 94.69%) Kazakhstanis have received an electronic health passport that stores their medical history and is available to the polyclinic, ambulance, hospital. Passport data is updated on a weekly basis with data from all medical institutions.

This study is aimed at the development of a software package for the diagnosis of clinical and hematological syndromes for an electronic passport, which allows automating the process of supporting clinical decision-making based on differential diagnosis algorithms and models of intelligent analysis of medical data. The aim of the study is to develop a software package of an electronic health passport based on information technology of differential diagnosis and computational and analytical models of morphological classification of clinical and hematological syndromes (CHS). The scientific foundations of the information technology for monitoring clinical and hematological syndromes integrate mathematical models for evaluating, pre-processing and ensemble methods of machine learning of biochemical blood test indicators, as well as methods that improve the efficiency of business processes for remote monitoring of hemoglobin synthesis disorders as a result of iron deficiency developing against the background of various pathological (physiological) conditions and manifested by signs of anemia and sideropenia.

The scientific novelty of the study lies in the fact that for the first time the technology of supporting clinical decision-making will be developed for differential diagnosis of CHS based on the use of ensemble machine learning methods. The relevance of the study is enhanced due to the

importance of applying the results of such studies for the East Kazakhstan region, since there is an increased level of radiation and exposure to chemicals in this region, which is why the prevalence of anemia among the population is increased.

Literature review

According to the classification of the World Health Organization (WHO), anemia is defined as a decrease in hemoglobin levels less than 120 g/l in women and less than 130 g/l in men. If we consider the threshold value proposed by WHO, it turns out that 22% of healthy people would be anemic [1].

Anemia is also widespread in Kazakhstan, as evidenced by the results of the National Nutrition Surveys conducted by the Kazakh Academy of Nutrition. The results of the study showed that more than 40% of school-age children suffer from anemia. The prevalence of anemia is especially high (49,4%) among children aged 12–14 years, as well as among women of reproductive age (48,2%) and among children aged 6–59 months (47,4%). It turned out that in Kazakhstan, almost every third man (28,1%) also suffers from anemia. Based on the entire population, the prevalence of anemia is 41,9%. This means that 6,5 million people in Kazakhstan suffer from anemia [2].

The processes of creating an electronic medical record and an electronic health passport are devoted to the works of many researchers who prove its effectiveness in the field of healthcare for various segments of the population. According to the definition, a health passport is a trusted document confirming the status of the owner in relation to certain diseases. [3].

In most studies, a cross-sectional study on the use of the passport was conducted among older people and their medical staff. A lot of controversy arises about the content of the health passport. The authors of [4], based on the results of studies that involved both patients (58%) and medical staff (42%), state that for a health passport to be most effective, its content should include the preferred elements of all stakeholders. As the content of the health passport, the authors used information of a different nature: on the progression of the disease; on the time scale of the disease; on a brief description of the visit to the doctor; on the results of the blood test; on the physical condition of the patient; on the dosages of medications taken .

In work [5], an attempt was made to create a scientific and methodological model for collecting and evaluating data on the physical health of schoolchildren, mental health, and physical fitness. The data is presented in the form of a digital health passport. These data were used as the basis for the computer program "Health Passport", which allows the user to receive information about the general state of health and differentiated aspects of the health of school-age children.

There are quite a few electronic resources dedicated to the use of medical data in the form of digital medical passes, which have become relevant during the pandemic. [6], [7]

The authors of [8] have developed a health passport aimed at involving students in controlling their own health levels according to various criteria of physical and mental development.

Despite significant advances in the use of technology in clinical medicine, the state of medical records remains fragmented and uneven due to a lack of standardized methods, structures, and standards. One of the challenges facing health information systems around the world is figuring out how to share medical records with multiple stakeholders for a wide range of purposes without compromising privacy and integrity [9].

In the modern world, it has become necessary to create a digital model of the patient, using various measurements that accumulate over a long period of observation. This task requires a consistent solution of such sequential tasks as: collection of patient data in the framework of digital medical devices (digital doctor's tools); storage of all data in uniform standards, the possibility of anonymous access to this data for the development of new methods and technologies for the treatment/diagnosis of patients; creation of methods for automatic processing of patient medical data to identify disease biomarkers [10].

To date, there is a large amount of hematological medical data collected because of medical tests. A lot of research is being done to gain knowledge from this data using data mining techniques. For example, the authors of [11] conducted three experiments with the test blood dataset using three classifiers: a decision tree, rule induction, and a naive Bayesian approach. The results showed that the naive Bayes classifier has a greater ability to predict blood disease than the other two classifiers. The method, called "naive Bayes", even though it is based on very simple assumptions, turns out to be quite competitive in some cases compared to more complex algorithms. This is especially common in the problems of medical diagnostics [12].

Deep machine learning for CHS indicators will improve the efficiency of differential diagnosis and apply it to develop algorithmic and software for an intelligent clinical decision support system [13, 14].

Thus, machine learning technology becomes an important tool in the development of new high-tech and personalized approaches to managing and monitoring CHS [15].

Research methods

The main scientific issues of the research are related to the development of the technology of differential diagnosis of CHS for an electronic passport, integrating mathematical models of evaluation, pre-processing, and ensemble methods of machine learning of indicators, biochemical blood analysis.

The main hypothesis of the study is the development and application of the technology of differential diagnosis of CHS for an electronic passport, based on algorithms of ensemble machine learning methods, which will increase the efficiency of technological processes for remote monitoring of hemoglobin synthesis disorders because of iron deficiency developing against the background of various pathological (physiological) conditions and manifested by signs of anemia and sideropenia.

For the implementation of scientific research, technologies based on interdisciplinarity will be applied, since the application of machine learning methods (the field of information technology) for business processes of monitoring indicators of biochemical blood analysis (medical sciences) is provided. Thus, a comprehensive application of differential diagnosis technology and machine learning algorithms is being implemented to improve the effectiveness of remote monitoring of clinical and hematological syndromes. Thus, to achieve the goal, it is necessary to consistently implement the main four stages of the study.

Stage one. For a comprehensive analysis of the problem and modeling of technological processes in the subject area, an information model of business processes will be developed and a conceptual design of a software package for the diagnosis of clinical and hematological syndromes for an electronic health passport will be implemented. At this stage, a descriptive approach will be initially applied: an intensive information and patent search will be conducted, a comparative analysis of the advantages and disadvantages of existing modern solutions will be performed, the choice of informative features for machine learning will be justified, a detailed experimental plan will be thought out, a literary review will be written. Also at this first stage, special research methods will be developed and applied, which are presented in Figure 1.

Stage two. To create a computational software and hardware complex, a set of mathematical evaluation models and ensembles of methods for data mining of clinical research indicators will be developed to support medical decision-making in the differential diagnosis of clinical and hematological syndromes. Within the framework of this second research task, the scientific issues presented in Figure 2 will be solved.

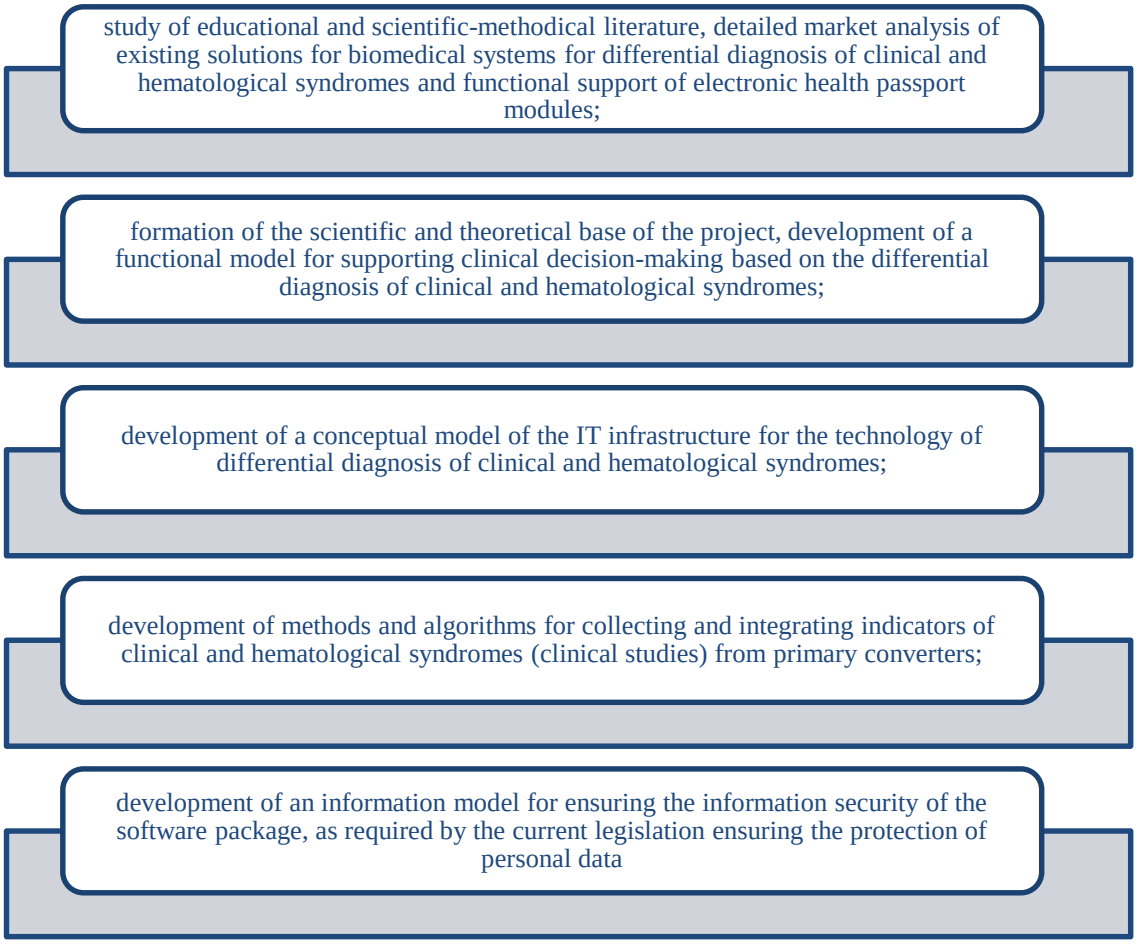

Figure 1. Scientific and technical issues of the first stage of the study

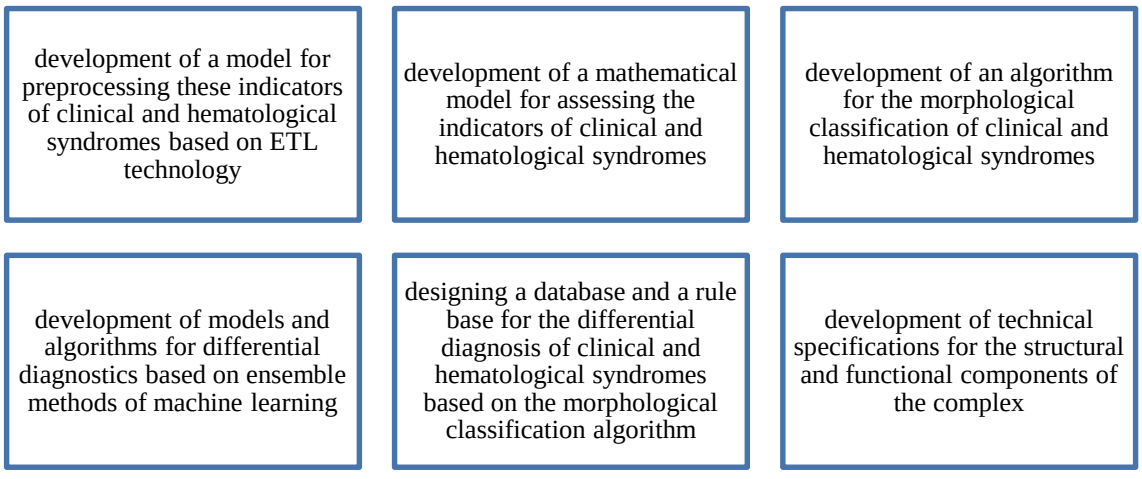

Figure 2. Scientific and technical issues of the second stage of the study

Models and algorithms of differential diagnosis will be based on such ensemble methods of machine learning as Backing and adaptive Boosting.

The Boosting algorithm implements the integration of machine learning algorithms (classifiers) to train the entire sample without splitting. When implementing the Boosting learning process, the algorithm allows you to change the weights of neurons, where analysis objects with incorrect classifications will have more weight and vice versa. Here, machine learning models aim to improve accuracy by analyzing the previous epoch (learning from errors).

The Backing algorithm allows you to implement 3 stages of classification: the stage of splitting the initial data into groups of bootstrap samples, building a classifier for each group, and implementing the final classifier.

Stage three. In accordance with the experimental research plan, architectural and software solutions for the implementation of the methods and algorithms developed (at the second stage of the study) will be developed, since the third stage is the development of an architectural and software solution for the diagnosis of clinical and hematological syndromes for an electronic health passport in a Web application that meets the requirements of information security and equipped with artificial intelligence functions and preventive analytics. Within the framework of this third research task, the scientific and technical issues presented in Figure 3 will be solved.

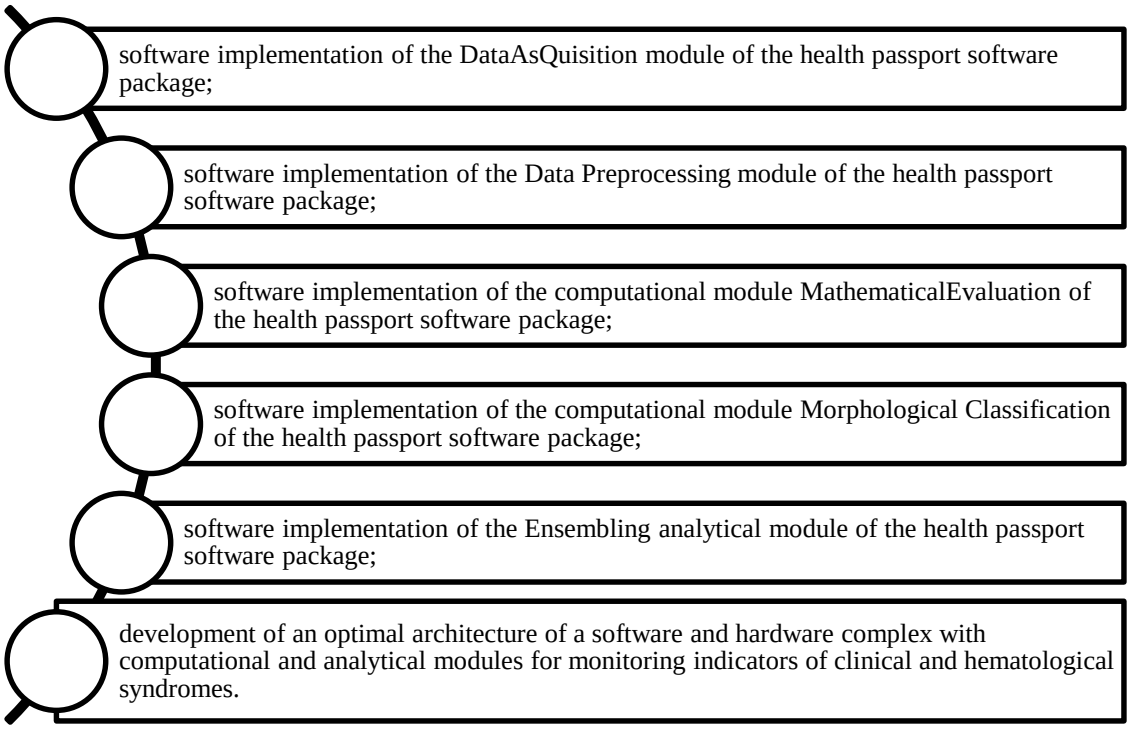

Figure 3. Scientific and technical issues of the third stage of the study

For the construction and software implementation of the proposed algorithms, the Python programming language will be used using the classes of the Scikit-learn library. Computational experiments with algorithms of ensemble machine learning methods will be based on the separation of the initial data on informative signs of human health and vital activity indicators into a training and test sample in the proportion of 70 to 30.

An optimal architectural model based on the use of components with the following properties:

reusability; substitutability; context independence; extensibility; encapsulation; independence.

Stage four. In accordance with the experimental research plan, experimental operation of the software and formalization of functionality based on experimental data will be carried out. Within the framework of this fourth research task, the scientific and technical issues presented in Figure 4 will be solved.

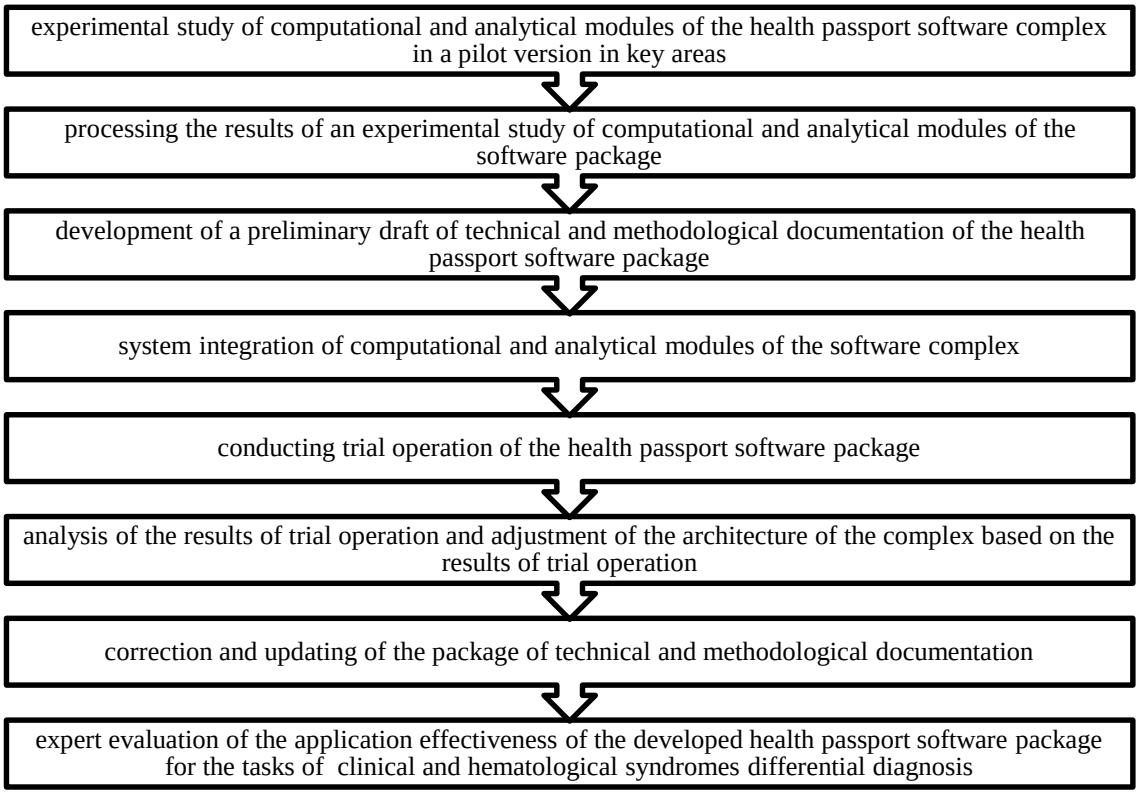

Figure 4. Scientific and technical issues of the fourth stage of the study

Thus, within the framework of this fourth task, computational and analytical modules will be tested and a preliminary draft of technical and methodological documentation of the software package will be developed.

The results of the study will make a significant contribution to the development of digital technologies for the transformation of administrative and medical processes of healthcare, aimed at improving the efficiency, quality, accessibility, and safety of medical care. In particular, the risk of medical errors will be reduced and the accuracy of classification of the data of differential diagnosis of anemia will be increased.

It is planned to carry out experimental work on approbation of practical developments in medical institutions of the East Kazakhstan region. The region has increased levels of radiation and exposure to chemicals, resulting in an increased prevalence of anemia among the population. Thus, the strategic task of the study is to increase the efficiency of diagnosing clinical and hematological syndromes by developing a comprehensive technology for differential diagnosis based on machine learning.

Conclusion

The scientific and technical effect of the study is to develop a new technology for monitoring

indicators of clinical and hematological syndromes based on differential diagnostics algorithms, integrating mathematical models for evaluation, pre-processing, and ensemble methods of machine learning for medical data. During the implementation of the project, new scientific results of medical data processing will be created by using ensemble methods of machine learning such as Backing and adaptive Boosting, designed to improve the accuracy of classifying clinical and hematological differential diagnosis data.

The creation of the software package will contribute to improving the level of the health protection and life safety of the Kazakhstan citizens , training specialists in the development of domestic medical systems and consolidation of medical data, strengthening digital culture, integration into existing medical information systems, introduction of innovative technologies, implementation of domestic platforms for professional development and retraining of specialists.

Acknowledgment

The work was carried out with the support of grant funding for scientific and (or) scientific and technical projects for 2023-2025 of the Ministry of Science and Higher Education of the Republic of Kazakhstan (grant №AP19679525).

Список литературы

1. Ракишева А.Г. Прогностическое значение анемии при хронической сердечной недостаточности. Диссертация на соискание степени доктора философии (PhD). Казахский национальный медицинский университет им. С. Д. Асфендиярова, 2018 г.
2. Рыспекова Н.Н. и др. Роль тяжелых металлов в развитии анемий (обзор литературы) // Вестник Казахского Национального медицинского университета. – 2019. – №. 3-2. – С. 46-51.
3. Health Passports: The Future Of Travel? Электронный ресурс. https://simpleflying.com/health-passports-future/
4. Johnson S. R. et al. Developing a pulmonary hypertension health passport to improve care: A qualitative study //Canadian Journal of Respiratory, Critical Care, and Sleep Medicine. – 2022. – Т. 6. – №. 3. – С. 193-198.
5. Vorobyov G.A., Chebotarev A.V., Panova I.P. The new scientific and methodological model of creating a digital health passport for monitoring physical and mental state and physical preparedness of comprehensive school students //Перспективы Науки и Образования Perspectives of Science and Education. – 2021. – Т. 54. – №. 6. – С. 522.
6. Digital Health Passes. What you need to know by bcd travel research & innovation. Электронный ресурс. https://docviewer.yandex.kz/view/0/?page=1&*=0lbR3y4LecrTD6xxQwkjrZC8cTR7InVy
7. Introduction to Digital Health Passports. Электронный ресурс. https://ts2.space/en/introduction-to-digital-health-passports/
8. Колос В.М., Петров Н.Я., Боярышев В.Н., Харитонович Г.С.. Индивидуальный паспорт здоровья студентов специального медицинского отделения. – Мн.: БГУИР, 2005. – 16 с.
9. Shelke P.M., Bagade J.V. Electronic Health Records: Need, Challenges, and Future Scope // Disruptive Developments in Biomedical Applications. – CRC Press. – С. 259-274.
10. Карпов О.Э., Храмов А.Е. Информационные технологии, вычислительные системы и искусственный интеллект в медицине. – М.: ДПК Пресс, 2022. – 480 с.
11. Dong N. et al. Medication Regularity of Traditional Chinese Medicine in the Treatment of Aplastic Anemia Based on Data Mining // Evidence-Based Complementary and Alternative Medicine. – 2022. – Т. 2022.
12. Червоненкис А.Я. Компьютерный анализ данных. Лекции Школы анализа данных Яндекса. М.: Яндекс, 2009. – 260 с.
13. Williams B. M. et al. An artificial intelligence-based deep learning algorithm for the diagnosis of diabetic neuropathy using corneal confocal microscopy: a development and validation study // Diabetologia. – 2020. – Т. 63. – №. 2. – С. 419-430.
14. Пустозеров Е.А., Тероева Ю.А. Применение машинного обучения для решения медицинских

задач. Конспект лекций. СПбГЭТУ «ЛЭТИ», 2021 г.
15. Kim E. et al. Towards more accessible precision medicine: building a more transferable machine learning model to support prognostic decisions for micro-and macrovascular complications of type 2 diabetes mellitus // Journal of medical.

References

1. Rakisheva A.G. Prognosticheskoye znacheniye anemii pri khronicheskoy serdechnoy nedostatochnosti. Dissertatsiya na soiskaniye stepeni doktora filosofii (PhD). Kazakhskiy natsional'nyy meditsinskiy universitet im. S. D. Asfendiyarova, 2018 g.
2. Ryspekova N.N. i dr. Rol' tyazhelykh metallov v razvitii anemiy (obzor literatury) // Vestnik Kazakhskogo Natsional'nogo meditsinskogo universiteta. – 2019. – №. 3-2. – S. 46-51.
3. Health Passports: The Future Of Travel? Elektronnyy resurs. https://simpleflying.com/health-passports-future/
4. Johnson S.R. et al. Developing a pulmonary hypertension health passport to improve care: A qualitative study // Canadian Journal of Respiratory, Critical Care, and Sleep Medicine. – 2022. – T. 6. – №. 3. – S. 193-198.
5. Vorobyov G.A., Chebotarev A.V., Panova I.P. The new scientific and methodological model of creating a digital health passport for monitoring physical and mental state and physical preparedness of comprehensive school students //Perspektivy Nauki i Obrazovaniya Perspectives of Science and Education. – 2021. – T. 54. – №. 6. – S. 522.
6. Digital health passes. what you need to know by bcd travel research & innovation. Elektronnyy resurs. https://docviewer.yandex.kz/view/0/?page=1&*=0lbR3y4LecrTD6xxQwkjrZC8cTR7InVy
7. Introduction to Digital Health Passports. Elektronnyy resurs. https://ts2.space/en/introduction-to-digital-health-passports/
8. Kolos V.M., Petrov N.YA., Boyaryshev V.N., Kharitonovich G.S.. Individual'nyy pasport zdorov'ya studentov spetsial'nogo meditsinskogo otdeleniya. – Mn.: BGUIR, 2005. – 16 s.
9. Shelke P.M., Bagade J. V. Electronic Health Records: Need, Challenges, and Future Scope // Disruptive Developments in Biomedical Applications. – CRC Press. – S. 259-274.
10. Karpov O. E., Khramov A. Ye. Informatsionnyye tekhnologii, vychislitel'nyye sistemy i iskusstvennyy intellekt v meditsine. – M.: DPK Press, 2022. – 480 s.
11. Dong N. et al. Medication Regularity of Traditional Chinese Medicine in the Treatment of Aplastic Anemia Based on Data Mining //Evidence-Based Complementary and Alternative Medicine. – 2022. – T. 2022.
12. Chervonenkis A.YA. Komp'yuternyy analiz dannykh. Lektsii Shkoly analiza dannykh Yandeksa. – M.: Yandeks, 2009. – 260 s.
13. Williams B.M. et al. An artificial intelligence-based deep learning algorithm for the diagnosis of diabetic neuropathy using corneal confocal microscopy: a development and validation study // Diabetologia. – 2020. – T. 63. – №. 2. – S. 419-430.
14. Pustozerov Ye.A., Teroyeva YU.A. Primeneniye mashinnogo obucheniya dlya resheniya meditsinskikh zadach. Konspekt lektsiy. – SPbGETU «LETI», 2021 g.
15. Kim E. et al. Towards more accessible precision medicine: building a more transferable machine learning model to support prognostic decisions for micro-and macrovascular complications of type 2 diabetes mellitus // Journal of medical.-and macrovascular complications of type 2 diabetes mellitus //Journal of medical.