DOI: 10.37943/17BRIJ4866

Indira Uvaliyeva
PhD, Professor of the School of Digital Technology and Intelligent Systems
indirauvalieva@gmail.com, orcid.org/0000-0002-2117-5390
D. Serikbayev East Kazakhstan Technical University, Kazakhstan

Aigerim Ismukhamedova
PhD candidate of the School of Digital Technology and Intelligent Systems
aigerim.ismukhamedova1@gmail.com, orcid.org/0000-0002-4931-337X
D. Serikbayev East Kazakhstan Technical University, Kazakhstan

Saule Belginova
PhD, Associate Professor of the Department of Information Technology
sbelginova@gmail.com, orcid.org/0000-0002-7238-6016
University Turan, Kazakhstan

Aigul Shaikhanova
PhD, Professor of the Department of Information Security
aigul.shaikhanova@gmail.com, orcid.org/0000-0001-6006-4813
L.N. Gumilyov Eurasian National University, Kazakhstan

IMPLEMENTATION OF A BASE OF RULES FOR DIFFERENTIAL DIAGNOSIS OF CLINICAL AND HEMATOLOGICAL SYNDROMES BASED ON MORPHOLOGICAL CLASSIFICATION ALGORITHM

Abstract: The evolving landscape of modern medicine underscores the growing importance of automating diagnostic processes. This advancement is not merely a convenience but a necessity to harness the full potential of technological progress, aiming to elevate research and clinical outcomes to new heights. Among the innovative strides in this field, the development of diagnostic systems based on morphological classification algorithms stands out. Such systems, rooted in comprehensive rule bases for differential diagnosis, promise to revolutionize the way we approach complex medical conditions. This paper introduces a cutting-edge system that epitomizes this evolution. Designed to harness the power of data analysis, it paves the way for groundbreaking research opportunities. At the heart of this system is a sophisticated set of rules derived from a morphological classification algorithm. This foundation enables the system to perform automated diagnoses of a wide array of clinical and hematological syndromes with unprecedented accuracy. A notable application of this technology is its ability to diagnose anemia by analyzing six distinct blood parameters and further categorize the anemia type based on biochemical criteria. The implications of such diagnostic capabilities are profound. By enabling the systematic collection and analysis of statistical data, the system facilitates in-depth research into the prevalence of diseases across different demographic groups. It aids in identifying disease patterns and supports preventive medicine efforts, potentially shifting the paradigm from treatment to prevention. This study not only highlights the system's capacity for enhancing diagnostic precision but also emphasizes its role as a catalyst for medical research and the improvement of healthcare delivery. The integration of such technologies into the medical field promises to enhance the quality of care, streamline diagnostic processes, and open new avenues for medical research, ultimately contributing to the advancement of global health standards.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 17, MARCH 2024

Keywords: rule base; artificial intelligence; automated systems; data analysis; medical information systems; differential diagnosis; clinical and hematological syndromes; morphological classification.

Introduction (Literary review)

In recent decades, with the rapid advancement of medical technologies and information systems, there has been a notable increase in interest in developing effective methods for differential diagnosis in the realms of clinical and hematological syndromes. This heightened interest stems from the ongoing expansion of the spectrum of recognized hematological diseases and the continual enhancement of diagnostic methodologies. Given the escalating volume of diagnostic data, there is a need to establish dependable, swift, and precise systems for processing and analyzing such data. An encouraging direction in this field entails setting up a rule base for distinguishing diagnoses through the use of a morphological classification algorithm. Artificial intelligence, as described by Amisha et al. (2019), can play a pivotal role in medicine by offering advanced diagnostic tools and decision support systems that help in identifying subtle patterns in clinical data [1].

The field of hematology is continuously evolving, with an increasing amount of information emerging annually regarding various syndromes and their morphological characteristics. This situation necessitates accuracy and promptness in diagnosis, leading to the advancement of innovative computer techniques and systems. Passamonti et al. (2022) suggest that the future of research in hematology will likely integrate conventional studies with real-world data and artificial intelligence, enhancing the reliability and efficiency of diagnostics [2].

Current methods for differential diagnosis, primarily reliant on empirical data and the expertise of professionals, may not consistently meet the demands efficiently. Hence, the creation of automated systems capable of analyzing and categorizing extensive clinical data becomes imperative for enhancing the quality and availability of healthcare services. Salib et al. (2021) discuss the integration of the Scopio Labs x100 Scanner with artificial intelligence capabilities into routine clinical workflows, which exemplifies the practical application of AI in improving diagnostic accuracy in hematology [3].

As the volume and complexity of medical data increases, the utilization of information technologies, such as artificial intelligence (AI) and machine learning, to process this data becomes crucial. Kaestner (2020) highlights how AI is meeting the challenges of hematology by facilitating the processing of complex datasets and contributing to more refined diagnostic processes [4]. The researchers also point out that, despite the active development of AI, rule-based systems remain popular and promising. Moreover, these systems are easier and more cost-effective to implement in medical process administration systems, which are not always equipped with modern technologies. The accuracy of rule-based systems in medicine can reach 80% - 90%, and they can be easily customized by adding or removing rules and adjusting weights [5-6]. The authors discuss the need for accessibility, widespread use, and a systematic approach in implementing such systems, including in mobile medicine [7]. A rule-base, crucial for the performance of rule-based systems through its “IF-THEN” statements that store, manage, and manipulate data for modeling complex problems [8]. These mechanisms are particularly useful in systems utilizing biological knowledge, as they enable the incorporation of information that is challenging to describe formally without disrupting the existing knowledge in the model [9]. This importance is underscored in a study that emphasizes the necessity of establishing a rule base to address specific medical issues, particularly in the realm of advancing hematology and biochemical analyzers [10].

Anemia is prevalent in Kazakhstan, as indicated by the findings of the National Nutrition Survey conducted by the Kazakh Academy of Nutrition. The survey results revealed that over

DOI: 10.37943/17BRIJ4866

40% of school-age children are affected by anemia. The incidence of anemia is notably high among children aged 12–14 years (49.4%), women of reproductive age (48.2%), and children aged 6–59 months (47.4%). It was discovered that in Kazakhstan nearly every third man (28.1%) also experiences anemia. Ryspekova (2019) discusses the role of heavy metals in the development of anemias, suggesting that environmental factors could be contributing to these high rates [11]. Across the entire population, the prevalence of anemia stands at 41.9%, indicating that 6.5 million individuals in Kazakhstan are affected by this condition. In context of using the national electronic healthcare system Sakko et al. (2023) provide additional factor by highlighting related health issues such as spontaneous pregnancy loss, which could be influenced by nutritional deficiencies prevalent across the country [12]. Bazarbaeva et al. (2021) further enrich our understanding by analyzing the morphofunctional and hematological characteristics of students from different regions of Kazakhstan, identifying regional disparities that might affect overall health outcomes [13]. Khozhayev et al. (2023) discuss the organization of hematological care in the Republic of Kazakhstan, providing insights into the healthcare system's response to managing diseases like anemia and ensuring proper care is available throughout the country [14]. These studies collectively illustrate a comprehensive picture of anemia's impact and the multifaceted approaches needed to address this widespread health concern in Kazakhstan.

In modern times, a significant volume of hematological medical data is collected through medical tests. Extensive research is underway to extract insights from this data utilizing data mining techniques. For instance, the authors conducted three experiments on a dataset 'test blood' employing three classifiers: decision tree, rule induction, and Naive Bayes [15]. The findings indicated that the Naive Bayes classifier exhibits superior predictive capabilities for blood disorders compared to the other two classifiers. Utilizing deep machine learning for the indicators of clinical and hematological syndromes (CGS) will significantly enhance the efficiency of differential diagnoses. This advancement is crucial for the development of algorithms and software that power intelligent clinical decision support systems. Consequently, formalizing the rule base emerges as a vital component in innovating new, high-tech, and personalized strategies for the management and monitoring of CGS.

Modern technologies allow the establishment of relationships between peripheral blood parameters in both normal and pathological conditions. The significance of automating laboratory processes is underscored in the research by Ralph Dadoun, emphasizing the value of automatically validating results against rule-based parameters to enhance laboratory productivity [16].

Special focus is placed on morphological features in hematology. Automated blood analyzers used for differential diagnosis frequently overlook morphological correlations, which are crucial for the diagnosis [17]. Chinese researchers have utilized AI-based tools to extract morphological features of lymphocytes for the analysis of chronic lymphocytic leukemia, demonstrating the potential for automation in this domain.

Indian scientists employ hemoglobin and hematocrit levels as input data for an expert system to identify anemia in children [18]. Researchers from Australia and Ghana highlight the efficacy of using neural networks in classifying anemia and evaluating hemoglobin levels [19]. Turkish scientists, utilizing four distinct artificial intelligence methods, achieved recognition accuracy rates of up to 85.6% in detecting anemia using hemograms and other patient information [20].

American researchers have developed an automated system for leukocyte classification based on morphological characteristics [10]. The utilization of deep neural networks has significantly improved the precision of platelet classification according to their morphological features, achieving accuracies of up to 98.6%. Al-qudah and Suen demonstrated that through

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 17, MARCH 2024

the use of enhanced incremental training techniques, the classification of blood cells in peripheral blood smears could be significantly improved, which supports these high accuracy levels [21].

Furthermore, Thinaharan and Thiagarasu describe a rule-based clinical decision support system for the healthcare industry, which integrates machine learning algorithms to further refine the accuracy and efficiency of medical diagnostics, including platelet classification [22]. Implementation of a base of rules in medical information systems promises to streamline the diagnostic workflow, making it more accessible and efficient for healthcare practitioners [23].

The innovative aspect of this research centers on creating and deploying a database model and software components for the «Health Passport» Web portal. This portal is distinctive because it utilizes a medical algorithm for the differential diagnosis of clinical and hematological syndromes.

At the national level, the project's scientific innovation is underscored by its pioneering status in the Republic of Kazakhstan; it marks the first development of a Web portal, «Health Passport,» dedicated to diagnosing clinical and hematological syndromes. This portal is remarkable for implementing a three-stage differential diagnosis process, which is grounded in a morphological classification rule base:

Stage 1 involves establishing the presence and severity of the syndromes;

Stage 2 determines their nature according to the morphological classification;

Stage 3 confirms and clarifies their nature using clinical biochemistry and enzyme immunoassay methods.

The global scientific novelty of the project stems from a rigorously scientific selection of diagnostic methods for clinical and hematological syndromes, anchored in the principles of a medical algorithm for differential diagnosis. The project's outcomes are poised to significantly advance the technological processes involved in the differential diagnosis of clinical and hematological syndromes. Furthermore, it will contribute to the monitoring of disrupted hemoglobin synthesis due to iron deficiency, a condition arising from various pathological (physiological) states and characterized by signs of anemia and sideropenia.

The core ambition of this investigation is to tackle a pressing challenge in the realm of clinical hematology - the enhancement of diagnostic methodologies for blood-related illnesses through the adoption of advanced computer technologies.

Rules for differential diagnosis of clinical and hematological syndromes

The diagnostic system for clinical and hematological syndromes relies on a morphological classification algorithm that facilitates disease identification, information storage for individual patients, official diagnosis formulation, and subsequent treatment selection [24]. This information system enables an initial patient diagnosis, which is segmented into three stages, each guided by a specific set of rules.

The first diagnostic stage involves determining the existence and intensity of the clinical and hematological syndrome [25]. This is based on the evaluation of the following six rules:

1) Rule for determining the level of pathology based on hemoglobin m_{HGB} (1):

$$m_{\mathrm{HGB}} = \begin{cases} 0, & \mathrm{HGB} \geq 125 \\ \frac{125-\mathrm{HGB}}{125-115}, & 115 < \mathrm{HGB} < 125 \\ 1, & \mathrm{HGB} \leq 115 \end{cases} \quad (1)$$

where m_{HGB} represents the level of pathology in relation to hemoglobin, and HGB denotes the hemoglobin measurement from the patient's sample.

2) Rule for determining the level of pathology based on the quantity of red blood cells m_{RBC} (2):

DOI: 10.37943/17BRIJ4866

$$m_{\text{RBC}} = \begin{cases} 0, & \text{RBC} \geq 4{,}0 \times 10^{12}/l \\ \frac{4{,}0-\text{RBC}}{4{,}0-3{,}5}, & 3{,}5 \times 10^{12}/l < \text{RBC} < 4{,}0 \times 10^{12}/l \\ 1, & \text{RBC} \leq 3{,}5 \times 10^{12}/l \end{cases} \quad (5)$$

where m_{RBC} represents the level of pathology determined by the quantity of red blood cells, RBC indicates the result obtained from the patient's sample.

3) Rule for determining the level of pathology based on hematocrit m_{HCT} (3):

$$m_{\text{HCT}} = \begin{cases} 0, & \text{HCT} \geq 0{,}38 \\ \frac{0{,}38-\text{HCT}}{0{,}38-0{,}30}, & 0{,}3 < \text{HCT} < 0{,}38 \\ 1, & \text{HCT} \leq 0{,}3 \end{cases} \quad (3)$$

where m_{HCT} represents the level of pathology determined by hematocrit, and HCT indicates the result obtained from the patient's sample.

4) Rule for determining the level of pathology based on the average hemoglobin content in a red blood cell m_{MCH} (4):

$$m_{\text{MCH}} = \begin{cases} 1, & \text{MCH} \leq 18{,}5 \text{ pg} \\ \frac{27-\text{MCH}}{27-18{,}5}, & 18{,}5 \text{ pg} < \text{MCH} < 27 \text{ pg} \\ 0, & 27 \text{ pg} < \text{MCH} < 34 \text{ pg} \\ \frac{MCH-34}{36{,}4-34}, & 34 \text{ pg} < \text{MCH} < 36{,}4 \text{ pg} \\ 1, & \text{MCH} \geq 36{,}4 \text{ pg} \end{cases} \quad (4)$$

where m_{MCH} represents the level of pathology based on the average hemoglobin content in the erythrocyte, MCH indicates the result of the patient's sample.

5) Rule for determining the level of pathology based on the average volume of red blood cells m_{MCV} (5):

$$m_{\text{MCV}} = \begin{cases} 1, & \text{MCV} \leq 64\ fl \\ \frac{80-\text{MCV}}{80-64}, & 64\ fl < \text{MCV} < 80\ fl \\ 0, & 80\ fl < \text{MCV} < 95\ fl \\ \frac{\text{MCV}-95}{129-95}, & 95\ fl < \text{MCV} < 129\ fl \\ 1, & \text{MCV} \geq 129\ fl \end{cases} \quad (5)$$

where m_{MCV} represents the level of pathology determined by the average volume of erythrocytes, and MCV indicates the result obtained from the patient's sample.

6) Rule for determining the level of pathology based on the average concentration of hemoglobin in a red blood cell m_{MCHC} (6):

$$m_{\text{MCHC}} = \begin{cases} 0, & \text{MCHC} \geq 32\ g/\text{dl} \\ \frac{32-\text{MCHC}}{32-28}, & 28\ g/\text{dl} < \text{MCHC} < 32\ g/\text{dl} \\ 1, & \text{MCHC} \leq 28\ g/\text{dl} \end{cases} \quad (6)$$

where m_{MCHC} represents the level of pathology determined by the average concentration of hemoglobin in erythrocytes, $MCHC$ indicates the result obtained from the patient's sample.

7) The rule for determining the severity of anemia M (7):

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 17, MARCH 2024

$$M = \begin{cases} 0, & \text{M}CHC \leq 0{,}2 \\ \sum M, & 0{,}2 < \text{M}CHC < 0{,}5 \\ 1, & \text{M}CHC \geq 0{,}5 \end{cases} \quad (7)$$

where $\sum M$ represents the overall degree of association with the anemic syndrome (M) for all six indicators (8):

$$\sum M = m_{HGB} * 0{,}5 + m_{HCT} * 0{,}1 + m_{MCHC} * 0{,}1 + m_{MCH} * 0{,}1 + m_{MCV} * 0{,}1 + m_{RBC} * 0{,}1 \quad (8)$$

The second stage involves directly identifying the type of anemia based on morphological classification. The membership function is computed for each type.

1. Associated with microcytic anemias.

The rule for anemia to belong to the microcytic type (9):

$$M_{micro} = \begin{cases} 0, & M_{micro} < 0{,}5 \\ 1, & M_{micro} \geq 0{,}5 \end{cases} \quad (9)$$

where M_{micro} is the overall criterion for anemia's classification as microcytic (10):

$$M_{micro} = m_{MCH} * 0{,}5 + m_{MCV} * 0{,}5 \quad (10)$$

For this type of anemia, the rules for determining the level of pathology based on the average hemoglobin content in the erythrocyte m_{MCH} (11) and determining the level of pathology based on the average volume of the erythrocyte m_{MCV} (12):

$$m_{\text{MCH}} = \begin{cases} 0, & \text{MCH} \geq 27\ pg \\ \frac{27-\text{MCH}}{27-18{,}5}, & 18{,}5\ pg < \text{MCH} < 27\ pg \\ 1, & \text{MCH} \leq 18{,}5\ pg \end{cases} \quad (11)$$

$$m_{\text{MCV}} = \begin{cases} 0, & \text{MCV} \geq 80\ fl \\ \frac{80-\text{MCV}}{80-64}, & 64\ fl < \text{MCV} < 80\ fl \\ 1, & \text{MCV} \leq 64\ fl \end{cases} \quad (12)$$

2. Associated with macrocytic anemias.

The rule for anemia to be classified as macrocytic (13):

$$M_{macro} = \begin{cases} 0, & M_{macro} < 0{,}5 \\ 1, & M_{macro} \geq 0{,}5 \end{cases} \quad (13)$$

where M_{macro} is the overall criterion for anemia's classification as macrocytic (14):

$$M_{macro} = m_{MCH} * 0{,}5 + m_{MCV} * 0{,}5 \quad (14)$$

For this type of anemia, the rules for determining the level of pathology based on the average hemoglobin content in the erythrocyte m_{MCH} (15) and determining the level of pathology based on the average volume of the erythrocyte m_{MCV} (16):

$$m_{\text{MCH}} = \begin{cases} 0, & \text{MCH} \leq 34\ pg \\ \frac{MCH-34}{36{,}4-34}, & 34\ pg < \text{MCH} < 36{,}4\ pg \\ 1, & \text{MCH} \geq 36{,}4\ pg \end{cases} \quad (15)$$

$$m_{\text{MCV}} = \begin{cases} 0, & \text{MCV} \leq 95\ fl \\ \frac{\text{MCV}-95}{129-95}, & 95\ fl < \text{MCV} < 129\ fl \\ 1, & \text{MCV} \geq 129\ fl \end{cases} \quad (16)$$

DOI: 10.37943/17BRIJ4866

3. Associated with normocytic anemia.

The rule for anemia to be classified as normocytic (17):

$$M_{norm} = \begin{cases} 0, & M_{norm} < 0{,}5 \\ 1, & M_{norm} \geq 0{,}5 \end{cases} \tag{17}$$

where M_{norm} is the overall criterion for anemia's classification as normocytic (18):

$$M_{norm} = m_{MCH} * 0{,}5 + m_{MCV} * 0{,}5 \tag{18}$$

In this case, the rules for determining the level of pathology based on the average hemoglobin content in the erythrocyte m_{MCH} (19) and determining the level of pathology based on the average volume of the erythrocyte m_{MCV} (20):

$$m_{\mathrm{MCH}} = \begin{cases} 0, & \mathrm{MCH} \leq 18{,}5\ \mathrm{p}g \\ \frac{\mathrm{MCH}-18{,}5}{27-18{,}5}, & 18{,}5\ \mathrm{p}g < \mathrm{MCH} < 27\ \mathrm{p}g \\ 1, & 27\ \mathrm{p}g < \mathrm{MCH} < 34\ \mathrm{p}g \\ \frac{36{,}4\ -\ \mathrm{MCH}}{36{,}4-34}, & 34\ \mathrm{p}g < \mathrm{MCH} < 38\ \mathrm{p}g \\ 0, & \mathrm{MCH} \geq 38\ \mathrm{p}g \end{cases} \tag{19}$$

$$m_{\mathrm{MCV}} = \begin{cases} 0, & \mathrm{MCV} \leq 80\ fl \\ \frac{\mathrm{MCV}-80}{90-80}, & 80\ fl < \mathrm{MCV} < 90\ fl \\ 1, & 90\ fl < \mathrm{MCV} < 95\ fl \\ \frac{100-\mathrm{MCV}}{100-95}, & 95\ fl < \mathrm{MCV} < 100\ fl \\ 0, & \mathrm{MCV} \geq 100\ fl \end{cases} \tag{20}$$

The third stage is dedicated to confirming and specifying the type of anemia based on biochemical criteria, specifically ferritin (Fer) and vitamin B_{12}:

1) Rule for determining anemia in chronic diseases (normocytic) (21):

$$M_{ACD} = \begin{cases} 0, & Fer \leq 40\ mcg/\mathrm{l} \\ \frac{Fer-40}{60-40}, & 40\ mcg/\mathrm{l} < Fer < 60\ mcg/\mathrm{l} \\ 1, & Fer \geq 60\ mcg/\mathrm{l} \end{cases} \tag{21}$$

where M_{ACD} represents the overall severity of pathology for anemia in chronic diseases, and Fer denotes the ferritin value obtained from the patient's sample results.

2) Rule for identifying iron deficiency anemia (microcytic and normocytic) (22):

$$M_{IDA} = \begin{cases} 0, & Fer \geq 40\ mcg/\mathrm{l} \\ \frac{40\ -Fer}{40-20}, & 20\ mcg/\mathrm{l} < Fer < 40\ mcg/\mathrm{l} \\ 1, & Fer \leq 20\ mcg/\mathrm{l} \end{cases} \tag{22}$$

where M_{IDA} represents the overall severity of pathology for iron deficiency anemia, and Fer denotes the ferritin value obtained from the patient's sample results.

3) Rule for identifying B_{12} deficiency anemia (macrocytic) (23):

$$M_{B12} = \begin{cases} 0, & B12 \geq 400\ \mathrm{n}g/\mathrm{m}l \\ \frac{400\ -B12}{400-100}, & 100\ \mathrm{n}g/\mathrm{m}l < B12 < 400\ \mathrm{n}g/\mathrm{m}l \\ 1, & B12 \leq 100\ \mathrm{n}g/\mathrm{m}l \end{cases} \tag{23}$$

where M_{B12} represents the overall severity of pathology for B_{12}-deficiency anemia, and B_{12} B12 indicates the value of vitamin B_{12} from the patient's test results.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 17, MARCH 2024

Therefore, the calculation stages presented above constitute the rule base of the developed system for detecting clinical and hematological syndromes, in particular anemia.

The developed rule bases will provide methodological (algorithmic) support to the computational and analytical modules of the "Health Passport" web portal's diagnostic software package for clinical and hematological syndromes (CGS). This software package, designed for the diagnosis of CGS within an electronic health passport framework, will automate clinical decision-making support. It accomplishes this through the use of differential diagnostic algorithms and models that facilitate the intelligent analysis of medical data.

Designing a database and web application based on a rule base.

The relational anemia diagnostic database was implemented in PostgreSQL using 'sequelize' for machine implementation. PostgreSQL is an open-source relational database management system (RDBMS) known for its extensibility and SQL compatibility. All background information, particularly data on hypoglycemia indicators and glycemic control monitoring, was obtained from MIMIC-III (Medical Information Mart for Intensive Care) [26]. This extensive and freely accessible database contains de-identified health data for more than forty thousand patients who were in intensive care units at Beth Israel Deaconess Medical Center from 2001 to 2021. The portal supports numerous scientific and analytical studies, including the improvement of clinical decision-making processes, epidemiology, and the development of electronic tools, all of which currently benefit from a wealth of experimental data.

The application that interacts with the database was developed in express.js, serving as the back-end for the front-end, which includes functionalities for adding and reading records in the database. The machine implementation model of the database is illustrated in Fig. 1.

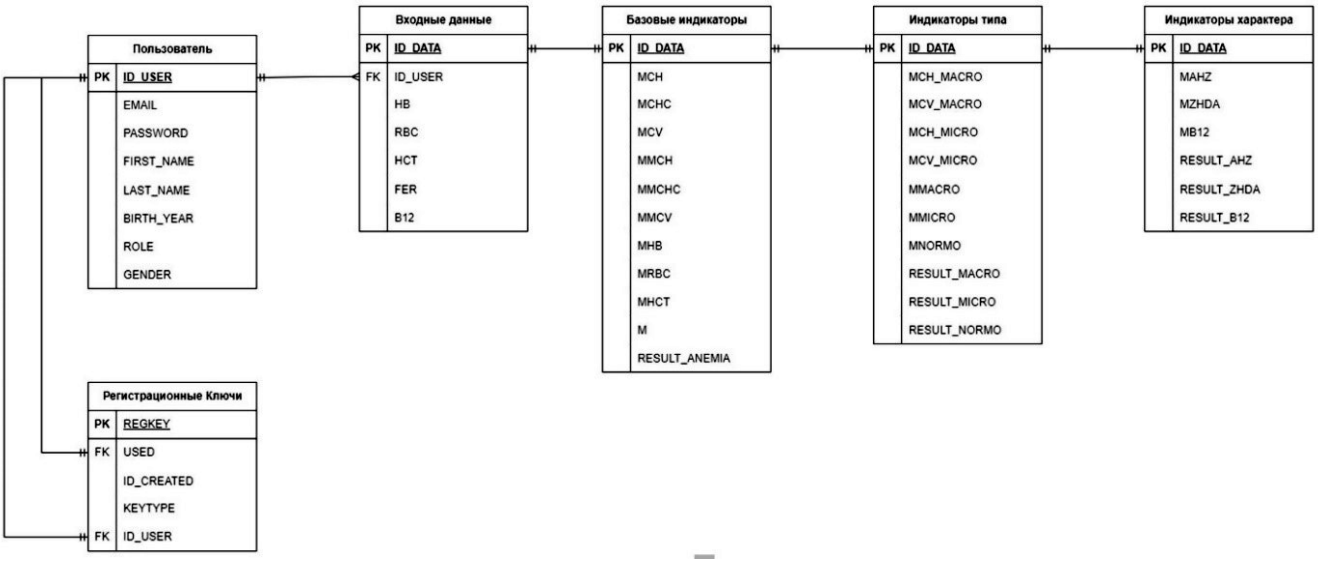

Figure 1. Physical model

Access to the system is provided via the "Health Passport" web portal. Each module within the system is accountable for a distinct stage. The sequential functioning of these modules is depicted in Fig. 2.

DOI: 10.37943/17BRIJ4866

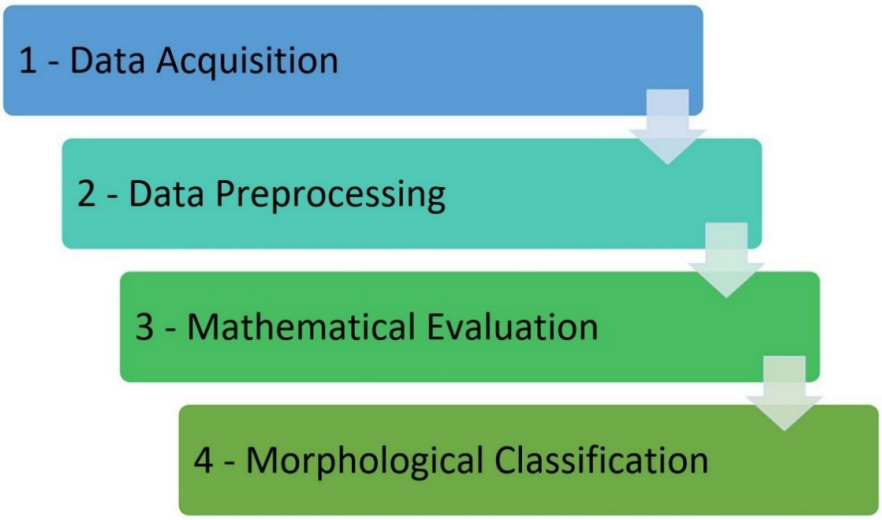

Figure 2. The sequence of processes for data management within the system.

The Health Passport web portal includes the following functional modules:

- Portal administration module;
- Data processing module (data acquisition, data preprocessing, mathematical evaluation, morphological classification);
- CDS diagnostic module;
- Diagnosis statistics module;
- «Patient Card» module;
- Patient diagnosis module.

The main window of the web application for implementing the rule base is depicted in the Fig. 3.

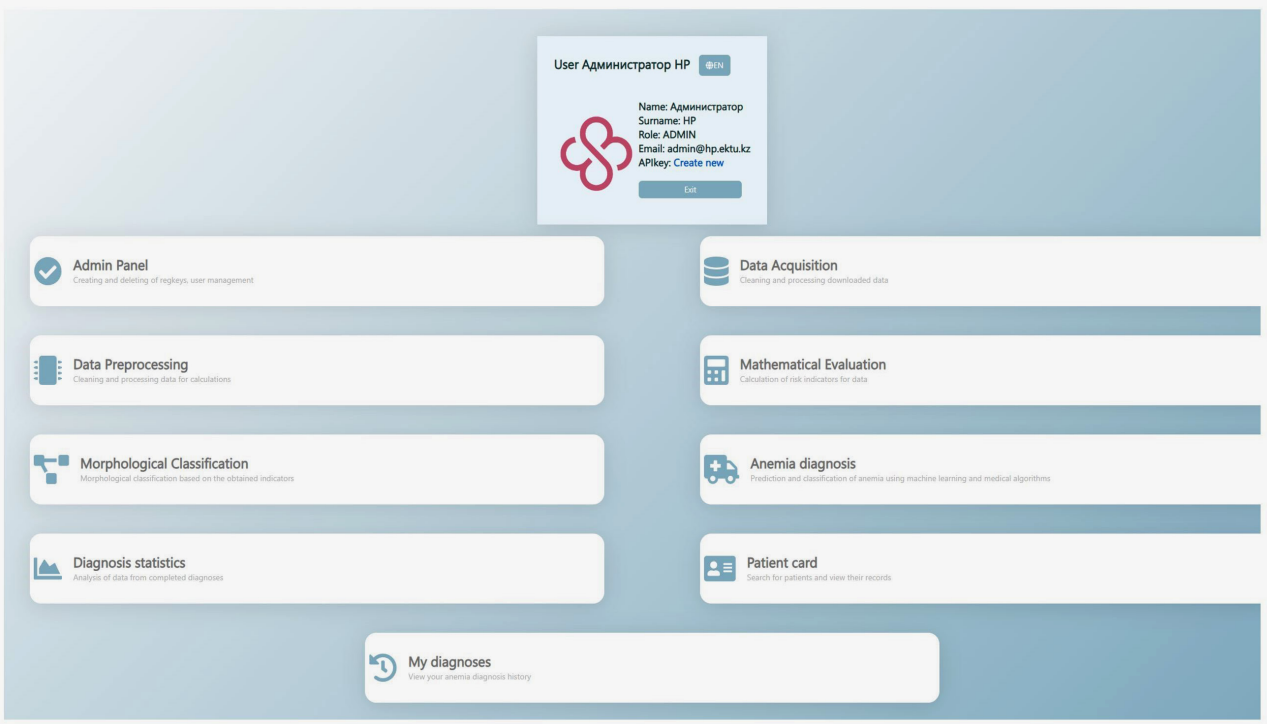

Figure 3. The main window of the web application for implementing the rule base

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 17, MARCH 2024

In the initial stage, data in Excel format is uploaded to the Health Passport. Table 1 displays a preview of the data in the "Data Collection" section, accessible to the administrator.

Table 1. Data Collection

id	gender	age	H8	RBC	HCT	Fer	B12	notes	Age_ group
7789	1	20	149	6,563876652	0,549396476	42	403	Gender 1 is female (1)	19-30
7790	1	30	159	6,25984252	0,450708661	62	390	You can leave your notes in this column	31-50
7791	0	40	90	4,186046512	0,298046512	17	268	0 is male	31-50
7792	0	35	149	9,3125	0,81484375				31-50
7793	1	25	147	6,681818182	0,664840909			Gender 1 is female	19-30
7794	1	65	116	5,201793722	0,387533632			0 is male	51+
7795	1	32	127	6,512820513	0,539912821	21	122	You can leave your notes in this column	31-50
7796	1	44	127	4,456140351	0,411301754	14	287		31-50
7789	1	20	149	6,563876652	0,549396476	42	403	Gender 1 is female	19-30

The partially processed data is subsequently utilized for mathematical calculations to ascertain the presence of anemia. This involves reducing the data dimensionality through feature selection or identifying the most significant attributes for analysis. Data preprocessing is a critical stage in data analytics and machine learning, encompassing a series of operations to cleanse and convert raw data for analytical purposes [27]. The pre-processing of loaded data incorporates the following procedures:

- Elimination of noise and anomalies. Utilizing statistical techniques to detect and remove outliers.
- Dealing with missing values. Addressing gaps by imputing mean, median, interpolation, or removing rows/columns with missing data.
- Data error correction. Rectifying evident errors like typos or inconsistent category usage.
- Normalization and standardization. Standardizing all numerical attributes to a uniform scale to enhance the convergence of machine learning algorithms.

The ultimate phase entails mathematical computations to identify the type of anemia based on morphological classification.

For each patient, a history of diagnoses and information regarding the indicators used for diagnosis is accessible; this section is illustrated in Fig. 4.

DOI: 10.37943/17BRIJ4866

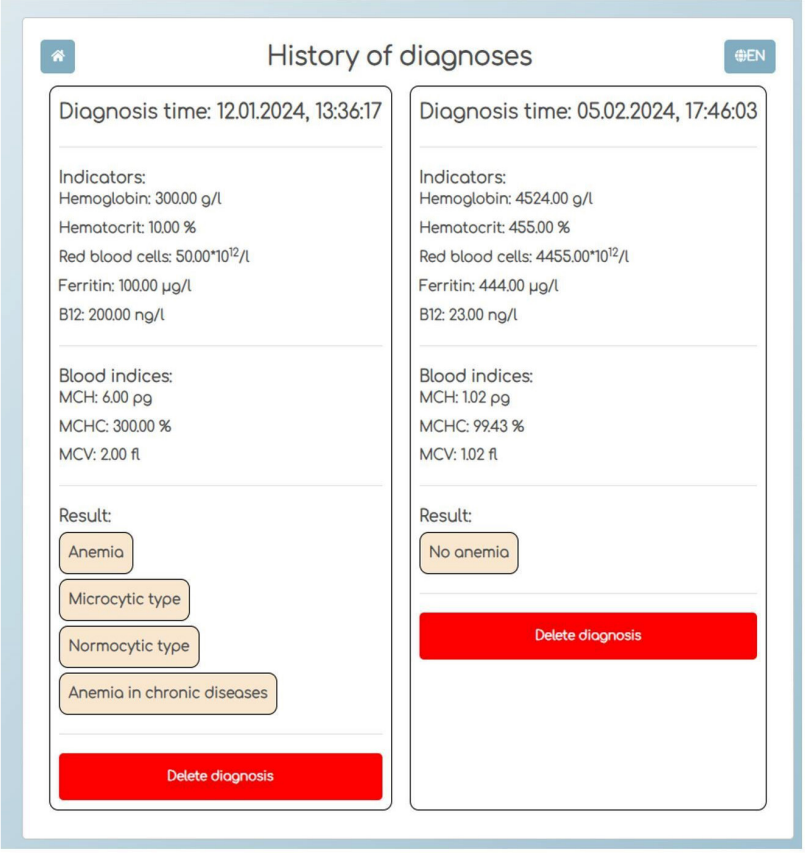

Figure 4. Patient diagnosis history window

The "Statistics" section aids to create visualizations of statistical patient data through charts and graphs. Graph construction can be based on variables like gender, anemia nature, anemia presence, anemia type, and year of birth. Filters can be applied to the selected data or custom filters can be created using SQL. Hence, generating statistics and utilizing a combination of NO-CODE technology requires only a basic understanding of database principles. For instance, Fig. 5 illustrates the outcome of statistical data in a chart format, with the condition "birth_year" = 1982 applied.

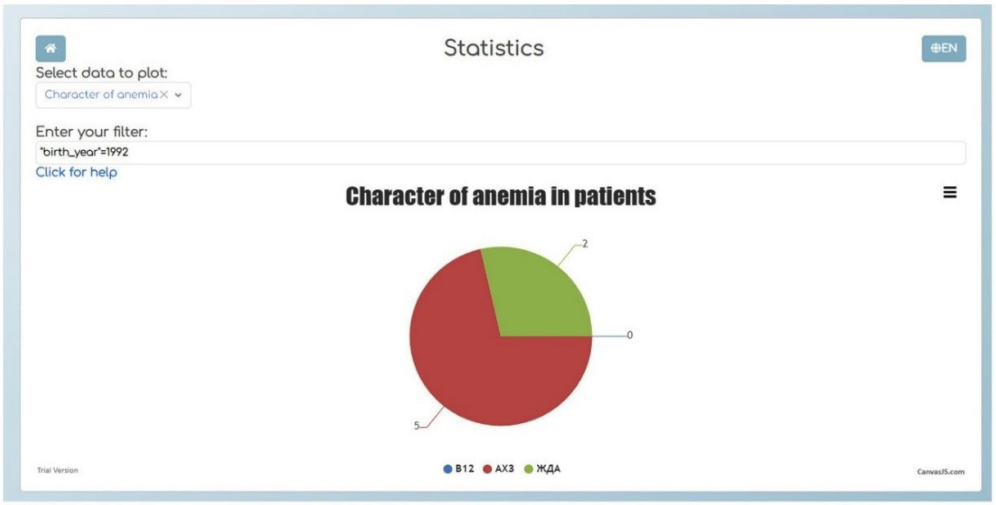

Figure 5. The sample of web-portal visualisation

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 17, MARCH 2024

Discussion

Comparison of the system proposed in this study with existing research demonstrates its superiority attributed to the utilization of a morphological classification algorithm, empowering a comprehensive analysis of the morphological features of blood cells. This highlights the potential of employing machine learning to augment our comprehension of biological and medical practices.

Nevertheless, the study is subject to several limitations. A primary concern is the restricted size and diversity of the training sample, potentially impacting the system's generalization capability. Subsequent investigations should contemplate enlarging the sample size to encompass a broader spectrum of clinical scenarios, thereby enhancing the system's accuracy and dependability.

Future research prospects entail the development and validation of algorithms for various hematologic conditions, alongside the integration of the system with electronic health passports to streamline the diagnostic process. Furthermore, it is imperative to explore the implications of making use of such systems on clinical decision-making and patient outcomes.

Conclusion

This study outlines the creation of an anemia diagnostic system that integrates a morphological classification algorithm with artificial intelligence and machine learning technologies. The findings reveal a notable enhancement in diagnostic precision and timeliness when compared to conventional approaches, emphasizing the value of incorporating advanced technologies into healthcare.

In conclusion, the research underscores the substantial potential of merging process rule base and machine learning in clinical diagnostics. The outcomes present promising opportunities for improving healthcare quality and availability, advocating for continued exploration and development of innovative diagnostic solutions.

Based on the advancement of clinical and hematological diagnostic methods, as well as their successful testing, implementation, and application, this study proposes a system designed to improve the process of diagnosing anemia. A system grounded in technological process rules is expected to strengthen the precision of the diagnostic process, with the ability to store collected data and results for subsequent research and the identification of new relationships between indicator combinations and disease detection. The developed system provides the effective analysis and classification of extensive clinical data, thereby improving the accuracy and efficiency of diagnosing hematological disorders.

The integration of modern technologies into clinical practice significantly contributes to improving the accessibility and quality of medical care. The findings of this study offer novel prospects for advancing and customizing diagnostic systems, emphasizing the significance of ongoing research in the application of artificial intelligence and machine learning in the medical domain.

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).

References

[1] Amisha, Malik, P., Pathania, M., & Rathaur, V. (2019). Overview of artificial intelligence in medicine. *Journal of Family Medicine and Primary Care*, *8*(7). https://doi.org/10.4103/jfmpc.jfmpc_440_19

[2] Passamonti, F., Corrao, G., Castellani, G., Mora, B., Maggioni, G., Gale, R.P., & Della Porta, M.G. (2022). The future of research in hematology: Integration of conventional studies with real-world data and artificial intelligence. *Blood Reviews*. https://doi.org/10.1016/j.blre.2021.100914
[3] Salib, C., Bhardwaj, S., Hussein, S., El Jamal, S., Petersen, B., Scigliano, E., Dembitzer, F.R., & Teruya-Feldstein, J. (2021). Digital AI in Hematology - Integration of the Scopio Labs x100 Scanner with Newly Implemented AI Capabilities into Routine Clinical Workflow. *Blood*, *138*(Supplement 1). https://doi.org/10.1182/blood-2021-148821
[4] Kaestner, L. (2020). Artificial intelligence meets hematology. *Transfusion and Apheresis Science*. https://doi.org/10.1016/j.transci.2020.102986
[5] Çelik Ertuğrul, D., Toygar, Ö., & Foroutan, N. (2021). A rule-based decision support system for aiding iron deficiency management. *Health Informatics Journal*, *27*(4). https://doi.org/10.1177/14604582211066054
[6] Kinney, E.L. (2022). Medical Expert Systems. https://journal.chestnet.org/article/S0012-3692(15)42851-X/fulltext
[7] Laosai, J., & Chamnongthai, K. (2018). Deep-Learning-Based Acute Leukemia Classification Using Imaging Flow Cytometry and Morphology. In *ISPACS 2018 - 2018 International Symposium on Intelligent Signal Processing and Communication Systems*. https://doi.org/10.1109/ISPACS.2018.8923175
[8] Yang, L.H., Ye, F.F., Liu, J., & Wang, Y.M. (2023). Belief rule-base expert system with multilayer tree structure for complex problems modeling. *Expert Systems with Applications*, *217*. https://doi.org/10.1016/j.eswa.2023.119567
[9] Troják, M., Šafránek, D., Pastva, S., & Brim, L. (2023). Rule-based modelling of biological systems using regulated rewriting. *BioSystems*, *225*. https://doi.org/10.1016/j.biosystems.2023.104843
[10] Rabbani, N., Kim, G. Y. E., Suarez, C. J., & Chen, J. H. (2022). Applications of machine learning in routine laboratory medicine: Current state and future directions. *Clinical Biochemistry*. https://doi.org/10.1016/j.clinbiochem.2022.02.011
[11] Ryspekova, N.N. (2019). Rol tyazhelyh metallov v razvitii anemij. *Vestnik Kazahskogo Nacionalnogo Medicinskogo Universiteta*, *3*(2), 46–51.
[12] Sakko, Y., Turesheva, A., Gaipov, A., Aimagambetova, G., Ukybassova, T., Marat, A., Kaldygulova, L., Amanzholkyzy, A., Nogay, A., Khamidullina, Z., Mussenov, Y., Almawi, W. Y., & Atageldiyeva, K. (2023). Epidemiology of spontaneous pregnancy loss in Kazakhstan: A national population-based cohort analysis during 2014–2019 using the national electronic healthcare system. *Acta Obstetricia et Gynecologica Scandinavica*, *102*(12). https://doi.org/10.1111/aogs.14669
[13] Bazarbaeva, S., Dinmukhamedova, A., Tleubergenova, G., Rakhimzhanova, Z., Sembekova, K., Karbayeva, S., & Kuandykova, E. (2021). Morphofunctional and hematological characteristics of health in students from the northern and southern regions of kazakhstan. *Open Access Macedonian Journal of Medical Sciences*, *9*. https://doi.org/10.3889/oamjms.2021.6434
[14] Khozhayev, A., Valiyev, R., Sultan, M., Tassybekov, Z., Zhumabaeva, M., Ryspekkyzy, A., Amangeldinov, Y., & Assylbek, R. (2023). Organization Of Hematological Care in the Republic of Kazakhstan. *Academics and Science Reviews Materials*, *4*.
[15] Dong, N., Zhang, X., Wu, D., Hu, Z., Liu, W., Deng, S., & Ye, B. (2022). Medication Regularity of Traditional Chinese Medicine in the Treatment of Aplastic Anemia Based on Data Mining. *Evidence-Based Complementary and Alternative Medicine*, *2022*. https://doi.org/10.1155/2022/1605359
[16] Syed, K., Sleeman, W. C., Ghosh, P., Nalluri, J. J., Kapoor, R., Hagan, M., & Palta, J. (2020). Artificial intelligence methods in computer-aided diagnostic tools and decision support analytics for clinical informatics. *Artificial Intelligence in Precision Health: From Concept to Applications*. https://doi.org/10.1016/B978-0-12-817133-2.00002-1
[17] Wang, Y., Liu, H., Wang, H., Wu, Y., Qiu, H., Qiao, C., Cao, L., Zhang, J., Li, J., Fan, L., & Wang, R. (2023). Enhancing morphological analysis of peripheral blood cells in chronic lymphocytic leukemia with an artificial intelligence-based tool. *Hematological Oncology*, *41*(S2). https://doi.org/10.1002/hon.3165_506
[18] Boadh, R., Chaudhary, K., Dahiya, M., Dogra, N., Rathee, S., Kumar, A., & Rajoria, Y. K. (2022). Analysis and investigation of fuzzy expert system for predicting the child anaemia. *Materials Today: Proceedings*, *56*. https://doi.org/10.1016/j.matpr.2022.01.094

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 17, MARCH 2024

[19] Appiahene, P., Chaturvedi, K., Asare, J.W., Donkoh, E.T., & Prasad, M. (2023). CP-AnemiC: A conjunctival pallor dataset and benchmark for anemia detection in children. *Medicine in Novel Technology and Devices*, *18*. https://doi.org/10.1016/j.medntd.2023.100244

[20] Karagül Yıldız, T., Yurtay, N., & Öneç, B. (2021). Classifying anemia types using artificial learning methods. *Engineering Science and Technology, an International Journal*, *24*(1). https://doi.org/10.1016/j.jestch.2020.12.003

[21] Al-qudah, R., & Suen, C.Y. (2021). Improving blood cells classification in peripheral blood smears using enhanced incremental training. *Computers in Biology and Medicine*, *131*. https://doi.org/10.1016/j.compbiomed.2021.104265

[22] Thinaharan, N., & Thiagarasu, V. (2019). A Rule Based Clinical Decision Support System for Healthcare Industry. *Indian Journal of Science and Technology*, *12*(12), 1–10.

[23] Uvaliyeva, I., Belginova, S., & Ismukhamedova, A. (2018). Development and implementation of the algorithm of differential diagnostics. In *Proceedings of the International Conference «Application of Information and Communication Technologies-AICT 2018»* (pp. 276–281). Almaty, Kazakhstan.

[24] Belginova, S., Uvaliyeva, I., & Rustamov, S. (2019). The application of data mining methods for the process of diagnosing diseases. *Journal of Theoretical and Applied Information Technology*, *97*(7).

[25] Uvaliyeva, I., Belginova, S., Rustamov, S., & Ismukhamedova, A. (2019). Algorithm Diagnosis of Anemia on the basis of the Method of the Synthesis of the Decisive Rules. In *13th IEEE International Conference on Application of Information and Communication Technologies, AICT 2019 – Proceedings*. https://doi.org/10.1109/AICT47866.2019.8981766

[26] Medical Information Mart for Intensive Care. URL: https://mimic.physionet.org/ (date of request: 10.03.2024).

[27] Uvaliyeva, I., Belginova, S., & Ismukhamedova, A. (2018). Informational and analytical system to diagnose anemia. In *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3234698.3234716