International Journal of Psychosocial Rehabilitation, Vol. 23, Issue 01, 2019
ISSN: 1475-7192

The Development of Expert Systems Using Artificial Intelligence

Rustam A. Burnashev, Ruslan G. Gabdrahmanov and
Arslan I. Enikeev

Abstract--- *This article presents a study related to the development of logical systems using artificial intelligence on the example of diagnosing leukemia. Data sets and logical rules that were used for the initial diagnosis of the disease were identified. In the process of learning, a neural network can reveal complex dependencies between the input data, as well as summarize the information obtained. This system includes a perceptron and consists of 3 hidden layers: 80 neurons (relu), 40 neurons (selu) and 1 neuron (sigmoid). For software development was used programming language Python.*

Keywords--- *Leukemia, Expert System, Python 3, PyQt5.*

I. INTRODUCTION

An artificial intelligence system is a system capable of performing creative or ambiguous functions that were previously considered exclusively human prerogatives, while AI researchers are free to use methods that are not applied by people if it is necessary to solve specific problems [Gloria., et al. 2014].

Machine learning systems are a kind of artificial intelligence systems whose principle of operation is learning in the process of solving many similar problems. To build such systems, the tools of mathematical statistics, numerical methods, optimization methods, probability theory, graph theory and various techniques for working with data in digital form are used.

Artificial neural network (neural network) - a mathematical model built in the image of the organization of biological neural networks - the nerve

cells of a living organism. This concept appeared when studying the processes occurring in the brain and trying to model them [Burnashev, Gubajdullin and Enikeev, 2018].

Neural networks usually consist of a set of elementary elements - neurons. Each such neuron is arranged quite simply, and deals only with the signals it receives and the signals it sends to other neurons. Neurons, in turn, are grouped into layers, the neurons of each layer in the classical model are usually connected according to the many-to-many scheme, i.e. each neuron receives information from the set of neurons in the previous layer and sends it to the set in the next. As a rule, a neural network consists of 1 input and output layers, as well as several “hidden” layers, which carry out the main part of the calculations. Such a system is already capable of solving rather complex tasks [Kamalov and Burnashev, 2017].

Rustam A. Burnashev, Institute of Computational Mathematics and Information Technologies, Kazan Federal University.
E-mail: burnashev@inbox.ru
Ruslan G. Gabdrahmanov, Institute of Computational Mathematics and Information Technologies, Kazan Federal University.
Arslan I. Enikeev, Institute of Computational Mathematics and Information Technologies, Kazan Federal University.
E-mail: a_eniki@inbox.ru

Received: 12 Nov 2018 | Revised: 22 Dec 2018 | Accepted: 06 Jan 2019

This means that after training, the neural network is able to return the correct result based on data that was missing from the data used for training, as well as noisy, incomplete, distorted data, but with less accuracy.

The medical knowledge base contains inference rules and information about human experience and knowledge in a certain subject area. In self-learning systems, the knowledge base also contains information resulting from the solution of previous problems [Burnashev, Yalkaev and Enikeev, 2017].

Leukemia is a term that unites numerous tumors of the hematopoietic system, arising from hematopoietic cells and affecting the bone marrow. The division of leukemia into two main groups: acute and chronic is determined by the structure of tumor cells: acute are leukemia's, the cell substrate of which is represented by blasts, and chronic - leukemia's, in which the bulk of tumor cells is differentiated and consists mainly of mature elements.

Today, the effectiveness of the treatment of acute leukemia throughout the world varies, depending on the form of the disease, from 30 to 90%. However, the fact that patients pass through suffering, as well as heavy and long-term treatment, should be taken into consideration. In a number of patients, it is necessary to use allogenic bone marrow transplantation in treatment programs. But if all the treatment is carried out on time, even in the most severe cases, the survival rate reaches 50%, and this is a lot, because before any leukemia was absolutely fatal disease.

In the course of the study, existing methods of diagnosing leukemia, known information about the nature of the course of the disease and about the changes that occur in the body of a sick person were studied. Also existing work in this direction, and the most popular types and architectures of artificial intelligence systems were reviewed. On the basis of the obtained data, the most practically applicable parameters were chosen that can signal the presence of the disease, promising in future to give a relatively high accuracy of diagnosis without the need to conduct complex and / or costly analyzes. Then, various approaches to the implementation of artificial intelligence systems were considered. This was aimed at selecting the type of system and architecture that can most correctly interpret the data obtained in the process of training and subsequent operation.

II. METHODS

To solve the tasks, the following sequence of actions was performed:

1. Formed requirements for expert system
2. Algorithm used to form requirements and Apyori library in Python
3. The study of existing methods of diagnosing leukemia.
4. Collect information about changes that occur with the patient's body.
5. Consideration of existing works in this area, various analogs and systems similar in purpose.
6. Structuring and analysis of the data.
7. Selection of the parameters used for the diagnosis, type and architecture of the future system based on the data obtained.
8. Creating a prototype knowledge base for training the system, filling it out by procedural data generation, since to obtain real data from free sources in sufficient quantities is difficult.
9. Design, training, testing of the prototype system.

Received: 12 Nov 2018 | Revised: 22 Dec 2018 | Accepted: 06 Jan 2019

To form requirements for the software being developed, an analysis was made of the area in which the expert systems will be used [Poole, Alan & Goebel, 1998; Burnashev and Enikeev, 2018] (Fig. 1, Fig. 2).

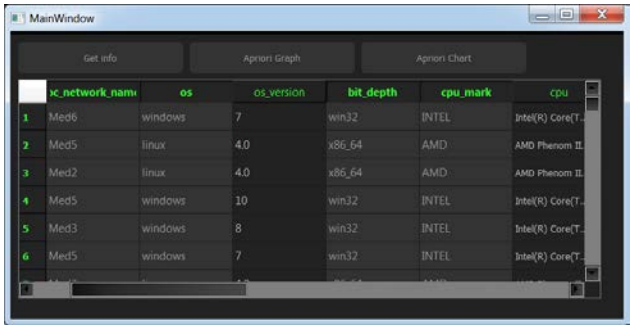

Fig. 1: Formation requirement interface

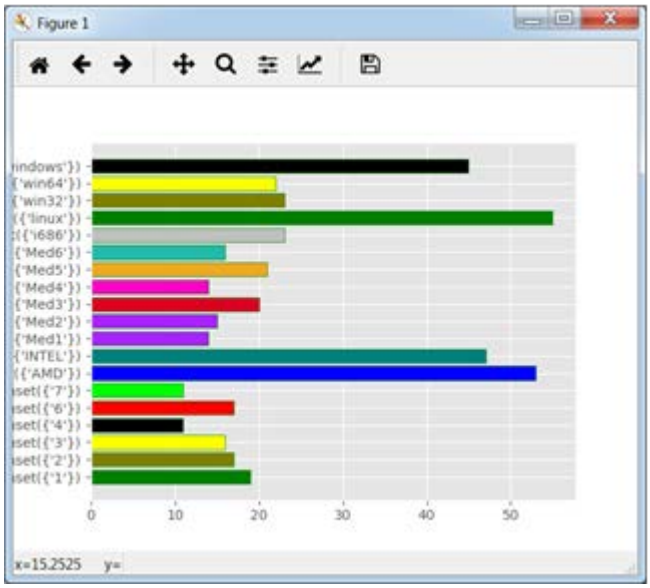

Fig. 2: The use of the matplotlib library for the visual presentation of the data

In the process of studying the existing methods of diagnosing leukemia, the literature on this topic and some scientific articles were studied.

The collection of information about changes that occur with the patient's body was carried out through the study of literature, articles, cases from practice.

Scientific works in the field of diagnosis of leukemia using artificial intelligence systems were considered.

The choice of the used parameters for the diagnosis was carried out on the basis of them:

- Reliability
- Accuracy
- Availability (data should be collected during standard medical prophylaxis - “physical examination”)
- Manufacturability (processing and interpretation complexity)

Received: 12 Nov 2018 | Revised: 22 Dec 2018 | Accepted: 06 Jan 2019

The choice of the type and architecture of the future system was based on the selected parameters for which the system should achieve:

- Theoretical and practical feasibility
- The greatest expected accuracy of diagnosis
- Acceptable from the point of view of practical operation speed and ease of use
- Available modern computer system requirements

The Python programming language was chosen for the development, training and testing of the system since it has an extensive set of libraries in the field of AI and data processing, as well as a fairly high computing speed, which will significantly speed up development. PostgreSQL DBMS is used for data storage[Lorier, 1991].

At the first stage, a general study of the existing methods of diagnosing leukemia was carried out. Puncture of the bone marrow is one of the main methods for clarifying the diagnosis and determining the method of treatment, but the method is not applicable for the indicated purposes due to the difficulty of carrying out for each patient.

Cytochemical blood test is used to diagnose types of acute leukemia. For these purposes it is not very promising, since it is not fully included in the medical examination procedure and covers only acute leukemia. Cytogenetic analysis and immunological diagnostics are also used to determine the type of leukemia, but are not included in the list of standard medical examination procedures and therefore are not suitable for these purposes.

Basing on the abovementioned, the most promising basic set of parameters for processing by the AI system are the results of a clinical blood test.

As already noted, the disease manifests itself with various symptoms, changes in the cellular and chemical composition of the blood, the physiological properties of the patient. Separately, it should be noted that almost all indicators of the clinical analysis of blood change, but all are noted in the medical literature [Gluzman, 2013].

III. RESULTS

To create a prototype neural network, it was decided to use the high-level libraries TensorFlow and Keras. They contain all the necessary tools to create different types of perceptrons, greatly speeding development. The PostgreSQL DBMS is used to store the generated data for learning, and the psycopg2 library is used to interact with it. The PyQt 5 library was used to develop the GUI shell.

During the design process, several prototypes of the neural network were created, which were integrated into the graphical interface and attached to the database. Thus, a prototype of an expert system was created based on the use of AI technologies that can assist specialists in diagnosing acute and chronic forms of leukemia.

The architecture (Fig. 5) of the most successful designed perceptron is as follows:

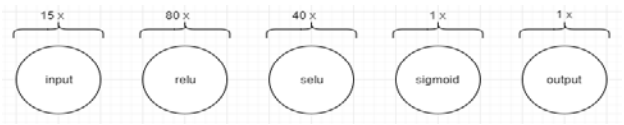

Fig. 3: Neural network architecture diagram

We describe the main points of the activation functions:

- relu (x, alpha = 0.0, max_value = None, threshold = 0.0) (Rectified Linear Unit), f (x) = max_value, if x> = max_value, f (x) = x, if threshold <= x <max_value and f (x) = alpha * (x - threshold) - otherwise;
- selu (x) (Scaled Exponential Linear Unit), scale * elu (x, alpha);
- elu (x, alpha = 1.0) (Exponential linear unit), x, if x> 0 and alpha * (exp (x) - 1), if x <0;

During the development of the system, the data was structured in the knowledge base with the specified architecture, which shows the following accuracy change in the learning process (Fig. 4):

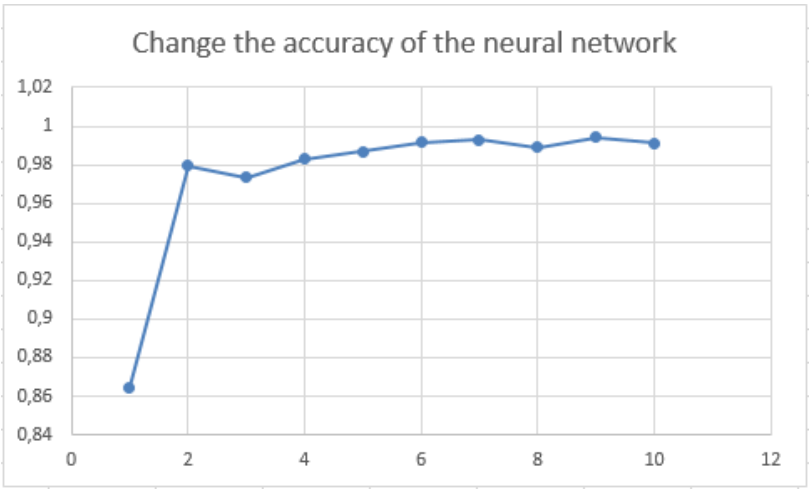

Fig. 4: Data structuring and learning process

As you can see, the neural network demonstrates relatively high accuracy (98-99%) with a fairly small drop in accuracy (Fig. 4) due to retraining of the generated data on close to reality, which suggests a theoretical possibility of successfully using similar systems in practice.

Basing on the data obtained, a user interface was developed in compliance with modern GUI requirements.

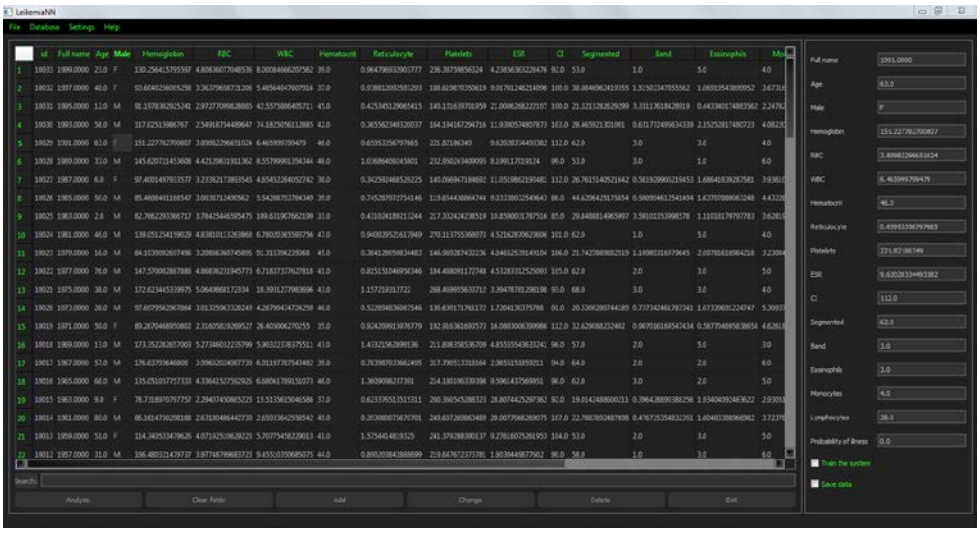

Fig. 5: The main form of the prototype system

Received: 12 Nov 2018 | Revised: 22 Dec 2018 | Accepted: 06 Jan 2019

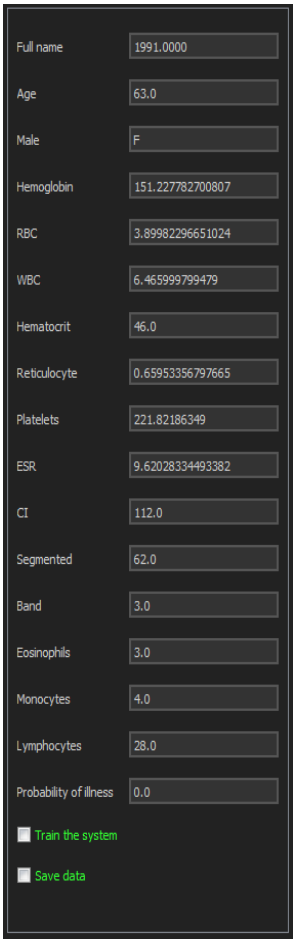

Fig. 6: Prediction Interface

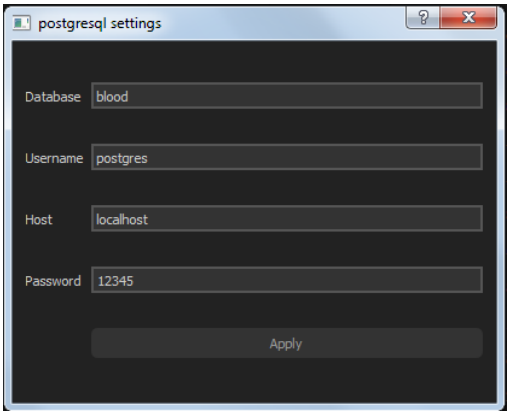

Fig. 7: Interface to connect to the database

IV. CONCLUSION

In the course of the research, the existing methods of diagnosing leukemia, the effect of the disease on the body, and the found analogues of information systems designed to diagnose leukemia were examined. Based on all the data collected, a theory was developed about the fundamental possibility of creating a system capable of conducting

primary diagnostics of leukemia based on data collected during the medical examination procedure.

Then a prototype neural network was created that can learn from the data provided (data from a clinical analysis of blood, age and sex) and predict the presence of the disease based on data of the same type that was used during training. The prototype was trained on close-to-reality generated data with a high randomness coefficient, and subsequently showed a high accuracy of diagnosis, which, although it cannot be considered a practical confirmation of the theory, but shows a theory that is possible and worthy of attention.

BIBLIOGRAPHY

[1] Requena, G., Sánchez, C., Corzo-Higueras, J.L., Reyes-Alvarado, S., Rivas-Ruiz, F., Vico, F., & Raglio, A. (2014). Melomics music medicine (M3) to lessen pain perception during pediatric prick test procedure. *Pediatric Allergy and Immunology*, *25*(7), 721-724.

[2] Burnashev, R.A., Gubajdullin, A.V., & Enikeev, A.I. (2018). Specialized Case Tools for the Development of Expert Systems. In *World Conference on Information Systems and Technologies*, 599-605.

[3] Kamalov, A.M., & Burnashev, R.A., (2017). Development of the system expert medexpert for differential disease diagnostics. *Astra Salvensis*, 55-64.

[4] Burnashev, R.A., Yalkaev, N.S., & Enikeev, A.I. (2017). *Journal of fundamental and applied sciences*, 9, 1403-1416.

[5] Poole, D., Mackworth, A., & Goebel, R. (1998). *Computational Intelligence: A Logical Approach.* New York: Oxford University Press

[6] Burnashev, R.A., & Enikeev, A.I. (2018**).** On the case study of life cycle. 10 (10), 1729-1734.

[7] Lorier J.L. (1991). *Artificial intelligence systems*. - M .: Mir, 1991. - 568 p. - 20 000 copies - ISBN 5-03-001408-X.

[8] Gluzman, D.F. (2013). Diagnosis of leukemia. *Atlas and practical guidance*, 2013.

Received: 12 Nov 2018 | Revised: 22 Dec 2018 | Accepted: 06 Jan 2019