METHOD FOR CONSTRUCTING A DECISION TREE FOR DIAGNOSING EPILEPSY IN CHILDREN

Maryam Midharara., Kuchkarova Nozimakhon Anvarovna

Abstract**:** A method and algorithm for constructing a decision tree for diagnosing epilepsy in children are proposed. The analysis and processing of medical data, consisting of electroencephalogram indicators, rhythms and provocative tests, which make it possible to determine the type of epilepsy in children, was carried out. The proposed algorithmic and software allows you to determine the type of epilepsy in children with a reliability and efficiency of 95%.

Keywords: expert system, electroencephalography, epilepsy, decision tree, method, algorithm, program.

Currently, the development of information technologies on the one hand and the emergence of the latest diagnostic and treatment technologies on the other hand require the use of artificial intelligence methods for processing and analyzing medical data and developing medical diagnostic systems. The use of expert systems improves the quality of medical diagnostics in clinical hospitals and medical institutions [1-4]. The study shows that the development of expert systems for the prevention, diagnosis of diseases and the improvement of the organization of neurological care for children and adolescents with epilepsy has not received sufficient attention, although it is of great interest to the medical community It should be noted that medical expert systems allow: to obtain more accurate results; use significant amounts of accumulated knowledge; improve the qualifications of medical workers; check own diagnostic decisions by doctors; turn to the computer for advice in difficult diagnostic cases [4-8]. Epilepsy is considered one of the most common chronic diseases of neurology and psychiatry. According to some scientific researchers, the prevalence in developed countries is 5.8% of patients per 1000 inhabitants, in developing countries this figure is - 10.3% of patients per 1000 inhabitants, and in rural - 15.4% of patients [9]. There are more than 50 million people with epilepsy in the world [10]. According to the WHO, epilepsy, despite the long history of the description of the disease and deep study, today is the second most disabled neurological disease [11]. It should be emphasized that the highest prevalence of epilepsy is observed in children. The frequency of epilepsy in childhood is, according to foreign data, 0.5% - 0.75% of the child population, and febrile seizures up to 5%. Approximately 1/3 of the pediatric population suffers from some form of epilepsy between the ages of 5 and 14 years, and in most adult patients, the initial onset of the disease occurs in childhood. One of the most important tasks in medicine of modern epileptology remains the search for the causes of the onset and development of epileptiform activity. The International League Against Epilepticism (ILAE) proposed for revision 6 categories of epilepsy classification according to etiology, according to which the genetic forms of epilepsy, structural, metabolic, immune, infectious and epilepsy with unknown etiology are distinguished [12].

The aim of the study is to develop a method, algorithm and software for constructing a decision tree for diagnosing epilepsy in children and adolescents.

Methods for analysis and processing of the patients' data.

Electroencephalography (EEG) is one of the simplest and most widely used examination methods for diagnosing epilepsy. Electroencephalography provides important information about the state of the bioelectrical activity of the brain and the form of an epileptic seizure, it is a mandatory research method for diagnosis, and also allows you to analyze the dynamics of the disease, the effectiveness of therapy, and when clinical remission is achieved, it determines the moment and rate of withdrawal of antiepileptic therapy. Electroencephalography has shown itself to be a highly informative, non-invasive and accessible diagnostic method [13-17].

Analysis and processing of the EEG of patients at risk for epileptic and acute symptomatic seizures is crucial for making important therapeutic decisions [18-21]. To solve these issues, a method for constructing a decision tree for determining epilepsy in children is proposed, which is described by the descriptor E. Each attribute of the descriptor determines a certain level of the decision tree. The attributes of the descriptor will be denoted with a set of indicators X, electroencephalogram rhythms R, and provocative samples P of patients. To solve these issues, a method for constructing a decision tree for determining epilepsy in children is proposed, which is described by the descriptor E. Each attribute of the descriptor determines a certain level of the decision tree. The attributes of the descriptor will be denoted with a set of indicators X, electroencephalogram rhythms R and provocative samples P of patients.

$$E = E\,(X,\ R,\ P),\ E_n \subset E, \quad (1)$$

where $n = 1 \div 4$.

Each descriptor (1) corresponds to a certain type of disease: febrile convulsions in children - E_1, benign rolandic epilepsy - E_2, juvenile myoclonic epilepsy - E_3, Lennox-Gastaut syndrome - E_4.. It should be noted that a certain disease E_n is determined by some subsets of indicators $X_p \subset X$,, where $p=0\div5$, electroencephalogram rhythms $R_s \subset R$, where $s=0\div6$, and provoking tests $P_d \subset P$, where $d = 0\div5$. At the same time, for each disease, we obtain, according to the decision tree, the resulting descriptor(1.1) in the form

$$E = E_n\,(X_p,\ R_s,\ P_d), \quad (1.1)$$

each disease is determined by certain indicators X, rhythms R and provoking tests P.

The method of constructing a decision tree for determining epilepsy in children consists of a sequence of solving the following tasks (Fig. 1):

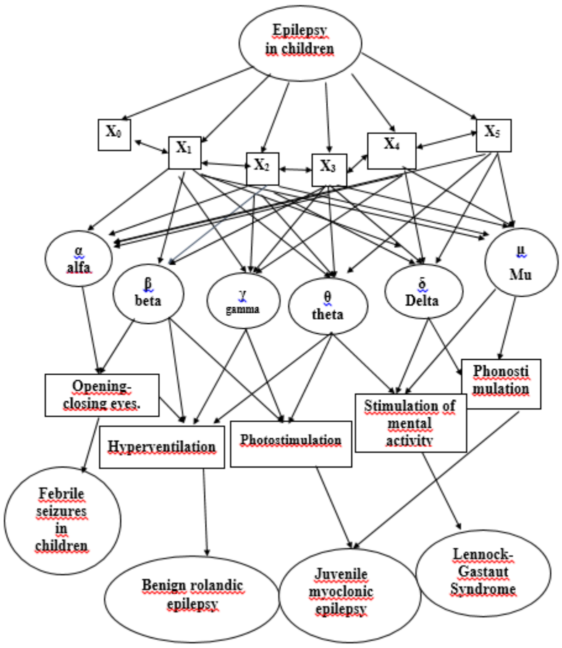

Fig.1. Decision tree for defining epilepsy in children

1.Patient data is fed to the root node of the tree.

2.Introduced EEG examination and basic EEG rhythms: alpha, beta, delta, theta, etc.

3.If, when examining EEG rhythms, pathological activity is determined (spikes, abrupt changes in the encephalogram), then an accurate diagnosis is made through provocative tests.

4.Paragraph 2 and paragraph 3 are repeated until a conclusion is reached, which represents the desired solution.

5.After all stages with epilepsy are determined.

To build a decision tree, first of all, we organize its root node consisting of six types of indicators:

1. Complete blood count determined by many parameters:

   $X_0 = \{X_{0,1}, X_{0,2}, X_{0,3},\ldots, X_{0,13}\}$, $X_{0,i} \in X_0$, where i= 1÷13; $X_{0,1}$ - saturation(SPO_2); $X_{0,2}$ – hemoglobin; $X_{0,3}$ – erythrocyte; $X_{0,4}$ - color index of blood (a blood test parameter expressing the relative content of hemoglobin in one erythrocyte, expressed in non-systemic units); $X_{0,5}$ – leukocyte; $X_{0,6}$ – bilirubin; $X_{0,7}$ – urea; $X_{0,8}$ – creatinine; $X_{0,9}$ –ESR; $X_{0,10}$ –platelets; $X_{0,11}$ – hematocrit; $X_{0,12}$ – lymphocytes; $X_{0,13}$ – monocytes.
2. Acquired external impact indicators are defined by multiple parameters $\mathbf{X_1 = \{X_{1,1}, X_{1,2}, X_{1,3},\ldots,X_{1,8}\}}$, $X_{1,j} \in X_1$, where j = 1÷8**:** $X_{1,1}$ - headache; $X_{1,2}$ - birth injury; $X_{1,3}$ - congenital anomalies of the brain; $X_{1,4}$ - chromosomal syndromes; $X_{1,5}$ - hereditary neurogenic syndromes; $X_{1,6}$ – infections; $X_{1,7}$ - a brain tumor; $X_{1,8}$ – hypoxia; $X_{1,9}$ – cranial nerve pathology.

3. Muscle spasm scores are determined by multiple parameters $\mathbf{X_2 =\{X_{2,1}, X_{2,2}, X_{2,3},\ldots,X_{2,7}\}}$, $X_{2,k} \in X_2$, where k = 1÷7: $X_{2,1}$ - hereditary metabolic defects; $X_{2\,2}$ - unnatural abrupt extension of the torso; $X_{2\,3}$ - muscle tension throughout the body; $X_{2\,4}$ - abrupt unreasonable shudders; $X_{2\,5}$ - subtle nodding of the head; $X_{2\,6}$ - legs bent, pulled up to the stomach; $X_{2\,7}$ - vascular pathology;

4. Indicators of brain pathology or perinatal lesions of the central nervous system are determined by many parameters $\mathbf{X_3 =\{X_{3,1}, X_{3,2}, X_{3,3},\ldots, X_{3,5}\}}$, $X_{3,m} \in X_3$, where m =1÷ 5 **:** $X_{3\,1}$ - behavioral deficits in the speech and facial areas; $X_{3\,2}$ - cerebral palsy; $X_{3\,3}$ - hydrocephalus (this is an accumulation of excess cerebrospinal fluid in the ventricles (hollow spaces) inside the brain. This leads to an increase in the size of the ventricles and excessive pressure on the brain [12]); $X_{3\,4}$ - encephalitis (then inflammation of the brain. [13]; $X_{3\,5}$ - prolonged somnolence (interventional epileptiform activity in the left paracentral region(cerebral palsy) is spontaneously registered);

5. The functional indicators of the human body are determined by many parameters $\mathbf{X_4 =\{X_{4,1}, X_{4,2}, X_{4,3}\}}$, $X_{4,n} \in X_4$, where n= 1÷3**:** $X_{4\,1}$ - Quantization Frequency (Hz); $X_{4\,2}$ - sleep deprivation (lack or complete absence of sleep); $X_{4\,3}$ - photostimulation (this is exposure to bright flashes of light directed into the patient's eyes);

6. Indicators of paroxysmal conditions are determined by many parameters $X_5 =\{X_{5,1}, X_{5,2}\}$, $X_{5,h} \in X_5$, where h=1÷4: $X_{5\,1}$ - paroxysmal activity (possible susceptibility to seizures); $X_{5\,2}$ - paroxysmal rapid sleep activity;

The second level of knowledge for the operation of making decisions in electroencephalogram rhythms $R=(R_\alpha, R_\beta, R_\theta, R_\delta, R_\gamma, R_\mu)$, $R_s \in R$, where $s = \alpha, \beta, \theta, \delta, \mu, \gamma$. Here R_α -alpha rhythm, R_β -

beta rhythm, R_θ -Theta rhythm, R_δ - delta rhythm, R_γ - Gamma rhythm, $R\mu$ - Mu rhythm. The third level of knowledge for the operation of the decision tree is provoking P tests. The results of these tests are extremely important for the attending physician in determining the correct diagnosis of the patient.

Provoking samples consist of many $P = \{P_1, P_2, P_3, P_4, P_5\}, P_d \in P$. Here P_1 - Opening-closing eyes.

P_2 – Hyperventilation, P_3 - Photostimulation, P_4 -Stimulation of mental activity, P_5 – Phonostimulation.

At the fourth level, various options for matching indicators, rhythms and stimulating trials for decision-making are analyzed, considered below from the listed 4 types of epilepsy: febrile convulsions in children, benign rolandic epilepsy, juvenile myoclonic epilepsy, Lennox-Gastaut syndrome.

Based on the proposed decision tree method, the stages of developing a medical expert system for diagnosing epilepsy in children are determined:

- organization of a decision tree management method;
- creating a database of patients with all the features and parameters required for the Dataset;
- diagnosis of the general condition of the patient upon admission to the clinic;
- processing and analysis of electroencephalograms;
- organization of communication between the levels of the decision tree;
- determination of the type of epilepsy;
- software for the implementation of algorithms;
- Evaluation and verification of the effectiveness of MES.

Database organization and software implementation of the algorithm.

The proposed MES makes it possible to predict epilepsy in children, it works according to a logic similar to that used by a doctor in everyday practice. To solve this problem, processing and analysis of the electroencephalogram is required. The development of the EEG database of children with epilepsy was carried out with the help of a neurophysiologist at the family clinic No. 45 in Tashkent, Uzbekistan.

Electroencephalography was performed in 370 patients suffering from various types of seizures under observation. Electroencephalograms were recorded on a 21-channel computer electroencephalograph NEURON-SPECTR-4 at the Family Clinic No. 45 in a screened room in the patient's sitting position.

Functional tests for hyperventilation and rhythmic photostimulation of various frequencies were used to stimulate biopotentials. Based on our research, a database was developed for the analysis of 370 patients (Fig. 2) based on 67 parameters in the form of an Excel table (Fig. 3). They were processed using Microsoft Excel computer programs for organizing Dataset. The dataset is divided into 4 parameter classes according to the 4 forms of epilepsy.

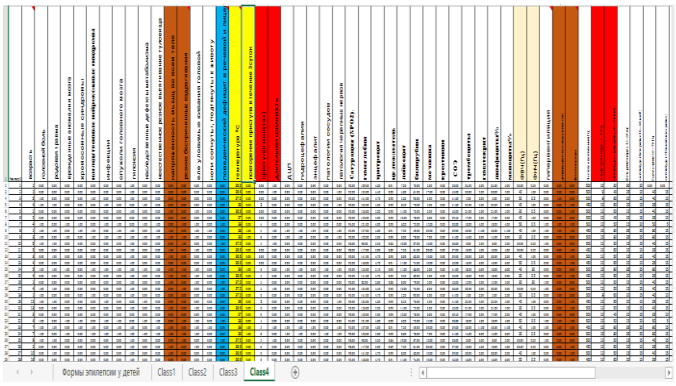

Fig.2. Database of patients.

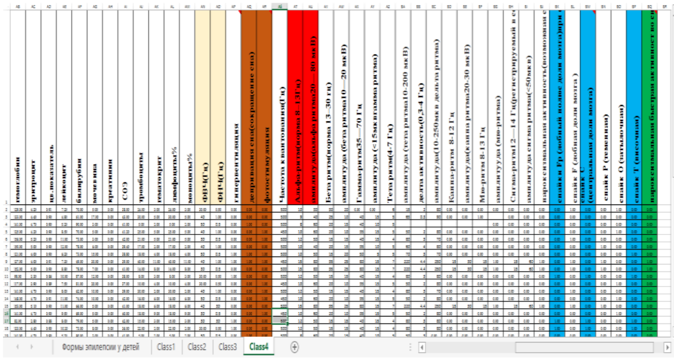

Fig.3. 2-part of the Database of patients.

The software implementation was carried out according to the Data Science CRISP-DM methodology and in the Python language.

The dialogue between the doctor and the expert system is implemented as a thin client (web application). For each visitor at the next visit, a unique session number is generated, by which the expert system distinguishes users. Using the data entry interface, the doctor compiles a list of symptoms and diagnoses that the patient has or assumes the doctor (with or without a positive coefficient of confidence). Excluded diagnoses and symptoms are also added (with a negative coefficient of confidence).

When developing the ES, I used the pandas library (for editing a dataset (dataset)) and numpy (calls the library for working with data in the form of a matrix), as well as matplotlib.pyplot (present data in graphical form). Colab allows you to use all the features of popular Python libraries for data analysis and visualization. For example, the cell below uses the numpy library to generate random data, as well as the matplotlib library to visualize it.

The implementation of the expert system was developed in the Python programming language using Google Colaboratory or Colab. This is another cloud service from Google Research. Also, it is an IDE that allows any user to write source code in an editor and run it from a browser. In particular, it supports the Python programming language and is focused on machine learning tasks, data analysis, educational projects, etc.

```
import pandas as pd #library to download dataset
import numpy as np # Causes traffic to work with data as a matrix
df = pd.read_excel('epilepsia_deti.xlsx') # Read db
df.isnull().sum() # checks if patients have null values in dataset
```

dfg = df.sort_values(['Vozrast','temperatura ⁰C'], ascending=False) # checks if patients have null values in dataset

import matplotlib.pyplot as plt # plot data.

After editing the library, we subtract all the parameters of the patient to the computer and highlight which parameters can be used to determine the type of epilepsy.

After executing the code in the age cell above on the page https://colab.research.google.com/drive, a scatter plot will appear that displays the temperature of patients (Fig. 4).

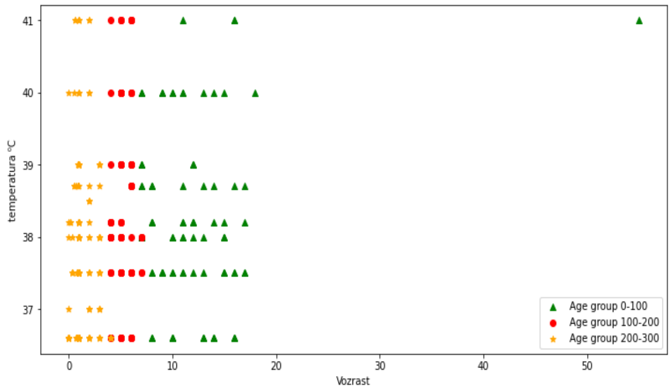

Fig.4. Diagram obtained by age and temperature of patients.

Based on the work of the software, the following results were obtained: 41 children correspond to the first type of epilepsy, 89 children to the second, 26 children to the third, 30 children to the fourth. At the same time, 186 patients were received, in which the diagnosis made by the developed program coincided with the diagnosis of the doctor, and the remaining 184 were healthy.

The effectiveness of the results is 95% when working with the medical data of 370 patients. If the number of patients is greater, then the efficiency will be increased accordingly.

CONCLUSION

A method for constructing a decision tree is proposed, which provides a treatment and diagnostic process in the diagnosis of epilepsy in children. Algorithms and their software implementation for determining the type of diseases have been developed and comparative assessment of the doctor's diagnosis of epilepsy by a physician and the results obtained on the basis of clinical and macroscopic examination data was made. A decision tree method is proposed that provides a therapeutic and diagnostic process for diagnosing epilepsy in children.

The developed software of the expert system can be used as an intellectual support for a practicing neuropathologist in interpreting the data of clinical and macroscopic examination of patients and making a preliminary diagnosis. When implementing MES programs in Python, the Data Science CRISP-DM methodology was chosen to implement the ES programs in the Python language.

The proposed algorithmic and software for determining epilepsy in children determines the type of disease with 95% reliability and efficiency. The software is created in Python, based on the developed dataset of patients with their disease parameters. As a result of the study of the disease, 370 patients received 186 patients (the rest are healthy), the diagnoses coincided with the diagnoses of the doctor.

References

Kononenko W, “Machine learning for medical diagnosis: history, state of the art and perspective, Artificial Intelligence in Medicine”. Ljubjana 2001. P. 89-109..

Ramesh A. N. et al. ”Artificial intelligence in medicine,” Annals of The Royal College of Surgeons of England. – 2004. – T. 86. – №. 5. – C. 334.

Patel V. L. et al. ”The coming of age of artificial intelligence in medicine,” Artificial intelligence in medicine. – 2009. – T. 46. – №. 1. – C. 5-17.

Furmankiewicz M., Sołtysik-Piorunkiewicz A., Ziuziański P. “Artificial intelligence systems for knowledge management in e-health: the study of intelligent software agents,” Latest Trends on Systems:The Proceedings of 18th International Conference on Systems, Santorini Island, Greece. – 2014, – C. 551-556.

Kumar S., Kaur G, “Detection of heart diseases using fuzzy logic,” International Journal of Engineering Trends and Technology. – 2013. – T. 38. – №. 6. – C. 2694-2699.

ONUWA O. B. Fuzzy, “Expert System for Malaria Diagnosis,” ORIENTAL JOURNAL OF COMPUTER SCIENCE & TECHNOLOGY. - 2014. - №7(2). - C. 273-284.

Kaur R., Kaur A,”Hypertension diagnosis using fuzzy expert system,” International Journal of Engineering Research and Applications (IJERA) National Conference on Advances in Engineering and Technology, AET-29th March. – 2014.

Kirillov V., Gladyshev A., Demidchik E, ”Technology of Creation of an Expert System for Diagnosing Thyroid Pathology Based on a Set of Qualitative Signs of Cell Atypia,” Microscopy Research and Technique. 2010. №73. C. 1091–1100.

Bell G. S., Neligran A., Sander J. W, “An unknown quantity –the worldwide prevalence of epilepsy,” Journal of Epilepsia. 2014; 55 (7): 958-962.

GBD 2016 Epilepsy Collaborators (2019). Global, regional and national burden of epilepsy, 1990–2016: a systematic analysis for the Global Burden of Disease Study. Lancet Neurol. 18: 357–375.

WHO Epilepsy Factsheet. Updated February 2017. URL: http://www.who.int/mediacentre/factsheets/fs999/en/. Accessed: 03.02.2017. 8.

R.B. Azizova Abdullayeva N.N. Usmonalieva I.I. Neuroimmunological Characteristics of Idiopathic and Symptomatic Epilepsy in Accordance with the Clinical Course. Medico-Legal Update. An International journal Volume 20, Number 4 October-December 2020. P 1377-1383.

Wheless, James & Clarke, Dave & Arzimanoglou, Alexis & Carpenter, Daniel. (2008). Treatment of pediatric epilepsy: European expert opinion, 2007. Epileptic disorders : international epilepsy journal with videotape. 9. 353-412.

Sanei Saeid and Chambers J.A. EEG signal processing, Centre of Digital Signal Processing Cardiff University, UK – John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2007. – C. 289.

Adeli H., Zhou Z., Dadmehr N, ”Analysis of EEG records in an epileptic patient using wavelet transform,” J. Neurosci. Methods. 2003. Vol. 123, no. 1, P. 69- 87.

Khan Y.U., Gotman Y, “Wavelet based automatic seizure detection in intracerebral electroencephalogram,” Clin. Neurophysiol. 2003. Vol. 114, no. 4, P. 898-908.

Wavelet-crosscorrelation analysis: non-stationary analysis of neurophysiological signals, Mizuno-Matsumoto Y. [et al], Brain Topogr. 2005. Vol. 17, no. 4, P. 237-252.

Chen H., Nui H, "Detection of character wave in EEG by wavelet," J. Electronic Sci. Technol. 2004. Vol. 2, no. 2, P. 269-271.

D'Atellis C.E., Isaacson S.I., Sime R.O, "Detection of epileptic events in electroencephalograms using wavelet analysis," Ann. Biomed. Eng. 1997. Vol. 25, P. 286-293.

Senhadji L., Wendling F. "Epileptic transient detection: wavelets and timefrequency approaches," Neurophysiol. Clin. 2002. Vol. 32, no. 3, P. 175-192.

21. Fast wavelet transformation of the EEG / Schiff S.J. [et al] , 'Electroencephalography and Clinical Neurophysiology', 1994. Vol. 91, no. 6, P. 442- 455.