 https://doi.org/10.1051/bioconf/20248401008

The use of artificial intelligence to diagnose the disease

Elena Suprun[1*], *Vadim* Tynchenko[1,2], *Vladimir* Khramkov[1], *Georgy* Kovalev[1], and *Tatiana* Soloveva[1]

[1]Reshetnev Siberian State University of Science and Technology, 660037 Krasnoyarsk, Russia
[2]Bauman Moscow State Technical University, 105005 Moscow, Russia

Abstract. This article explores the use of artificial intelligence in the medical field for diagnosing a disease, namely the identification of factors that affect the presence of a brain tumor. Modern medical technologies are developing rapidly, and artificial intelligence is becoming an increasingly important tool to help doctors in accurate and timely diagnosis of various diseases. The article focuses on the application of learning methods such as decision trees, Kohonen maps and neural networks. The development and application of artificial intelligence in medicine provides a huge potential for improving the diagnosis of diseases and increasing the effectiveness of treatment, which contributes to improving the quality of life of patients. However, do not consider the need for ongoing scientific support, testing and regulation to ensure the safety and reliability of the application of artificial intelligence in medicine.

1 Introduction

The impact of artificial intelligence, namely machine learning, on various fields of activity is significant. With its help, a large amount of data is analyzed, self-learning occurs and the result is given. Based on such results obtained by finding patterns, he can draw conclusions, predict events or make decisions. Such systems are close to human intelligence and are not able to go beyond the functions inherent in them [1, 2].

Many medical organizations use neural networks in their work [3]. Their use allows to reduce the time spent on diagnostic procedures, and also provides the medical staff with information for making more accurate diagnoses [4]. The main tasks of artificial intelligence are to improve the efficiency of the healthcare system and reduce the burden and volume of the routine work of doctors, allowing them to concentrate on making accurate diagnoses. Among the main areas of artificial intelligence in medicine are:

1. Monitoring the patient's condition [5];
2. Automated analysis of medical images;
3. Automation of routine tasks;
4. Processing and analysis of big data [6, 7];
5. Assistance in making medical decisions [8].

* Corresponding author: anna.glinskaya.02@gmail.com

https://doi.org/10.1051/bioconf/20248401008

The aim of the work is to obtain machine learning models that will help in making decisions related to the diagnosis of a brain tumor.

The use of machine learning for decision making accelerates and improves the quality of decisions in the medical field. The constructed models will contribute to a more accurate diagnosis in patients [9].

2 Materials and method

Brain tumors are a group of various neoplasms that occur inside the skull. This occurs as a result of the process of rapid and uncontrolled division of cells that were previously healthy components of the brain tissue [10, 11]. Such neoplasms begin to lead an autonomous existence in the human body. They can be either malignant or benign.

Brain cancer is a rare, poorly understood, and often fatal disease [12]. At the same time, according to doctors, a characteristic feature of cancer patients is the extreme neglect of the disease, when the chances of a cure are much less than they could be [13].

There are many studies to identify factors influencing tumor development. Each author identifies different factors. For example, V. M. Gaidukov, when evaluating, says that the first signs of a brain tumor in men depend on the size of the tumor, its type and localization. Benign neoplasms develop slowly, without symptoms, and last for several years in a latent form with periodic exacerbations of the clinical picture [14]. V. A. Belchenko considers the development of brain cancer in children, which is often associated with gene abnormalities. For example, glioma (pilocytic astrocytoma) diagnosed in children is caused by von Recklinghausen syndrome in almost 50% of cases [15, 16].

However, these studies do not describe all possible relationships between the factors that determine the occurrence of a tumor. For this reason, it was decided to conduct this study to find out the factors that influence or do not influence the development of a brain tumor [17].

To simplify the work with data, the data analysis method is used, which allows automating the construction of analytical models, also called machine learning [18]. The method can adapt, learn, and predict new data. Data processing was carried out using the Kohonen map. The Kohonen map is a neural network that uses unsupervised learning and performs the task of visualization and clustering [19]. The initial data are multidimensional, therefore, for a better perception, they must be visualized, which is what Kohonen's maps are intended for. Decision Trees (DT) is a non-parametric supervised learning method used for classification and regression [20, 21]. The goal is to create a model that predicts the value of the target variable by learning simple decision rules derived from the characteristics of the data. Neural networks are computational structures that model simple biological processes similar to those occurring in the human brain. Neural networks are capable of adaptive learning [22].

3 Results

As the data under study, a dataset was taken containing a data set of features of a brain tumor, including five first-order features and eight texture features with a target level. First order features: type, variance, standard deviation, symmetry, kurtosis. Second order features: contrast, energy, ASM (angular momentum), entropy, homogeneity, dissimilarity, correlation, roughness.

Of the methods considered for analysis, the Decision Tree method became the most accurate. We will consider it in detail.

Before you start working with data, you need to normalize it. This is used to achieve better sampling conditions. The purpose of normalization is to bring data in different units of measurement to a single form that will allow them to be compared with each other [23, 24].

Analysis through a decision tree is based on finding patterns among the data and predicting the result. We leave the standard settings and display the decision tree in the form of a diagram, as well as the number of errors. The decision tree diagram and the number of errors are shown in Figures 1 and 2.

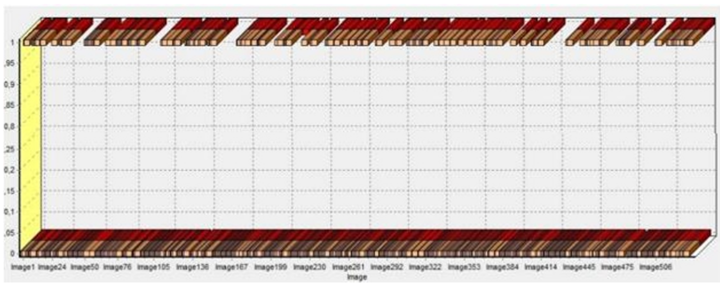

Fig. 1. Number of errors (scatterplot)

	Classified		
Actually	0	1	Total
0	378		378
1	3	152	155
Total	381	152	533

Fig. 2. Number of mistakes

From the results of the work, we see that we have only 7 errors out of 533 or 3.4%, which is a good result. From the correlation carried out during the analysis, it can be seen that the main significant factor for the decision tree is the presence of entropy. The correlation is shown in Figure 3.

Target attribute: type		
№	Attribute	Significance
1	Entropy	93,887
2	Asymmetry	3,876
3	Uniformity	2,237
4	Energy	0,000
5	ASM	0,000
6	Difference	0,000
7	Correlation	0,000
8	Dispersion	0,000
9	Average value	0,000
10	Standard deviation	0,000
11	Contrast	0,000
12	Excess	0,000

Fig. 3. Significance of correlation

The conditions on which the decision tree is based were also specified. The introduced conditions are shown in Figure 4.

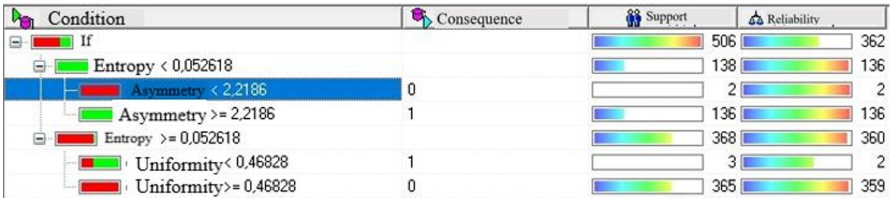

Fig. 4. Conditions for building a model

From the data obtained, it can be seen that when building the model, the main emphasis was placed on factors such as entropy, asymmetry, and homogeneity.

4 Discussion

Machine learning is a powerful tool for data analysis. Its main advantage lies in the ability to find patterns in various complex data sets. Traditional data analysis methods are often based on subjective assessments and predetermined metrics and do not fully capture the essence [25, 26].

Machine learning models are changeable. That is, they can be retrained and changed for subsequent data analysis [27-29]. In order to properly train the model, it is necessary to carefully process the target data [30]. It should also be noted that for individual datasets from different medical and other institutions, the models required for analysis cannot be universal. For each specific institution, for the identification of a specific diagnosis, individual decision-making models are built [31, 32].

5 Conclusion

The purpose of the work is to identify factors, the totality of which can be used to diagnose the presence of a brain tumor. Of the factors most accurately indicating the presence of this kind of diagnosis, the presence of entropy stands out. The resulting model can be used with other clinical and laboratory studies to improve the accuracy of the diagnosis. The presence of entropy in some cases can be a key identifier of the tumor process, which makes it a useful tool for specialists in various fields. An in-depth study of this factor may lead to the development of more effective methods for diagnosing and treating brain tumors, thereby improving patient outcomes and improving their quality of life.

Overall, these results represent an important contribution to the understanding of diagnostic factors in brain tumors and guide future research in this area.

References

1. B. V. Malozyomov. Micromachines, **14** (7)1288 (2023)
2. I. I. Bosikov, N. V. Martyushev, R. V. Klyuev, I. A. Savchenko, V. V. Kukartsev., V. A. Kukartsev & Y. A. Tynchenko, Fire, **6(3),** 95 (2023)
3. A. S. Mikhalev et al., Symmetry **14.10** (2022): 2036.
4. K. Moiseeva et al. *The Impact of Coal Generation on the Ecology of City Areas* //2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH). – IEEE, 1-6 (2023)

5. V. Kukartsev et al. *Analysis of Data in solving the problem of reducing the accident rate through the use of special means on public roads* //2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). – IEEE, 1-4 (2022)
6. T. Kireev et al. *Analysis of the Influence of Factors on Flight Delays in the United States Using the Construction of a Mathematical Model and Regression Analysis* //2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). – IEEE, 1-5 (2022)
7. V. Kukartsev et al. *Prototype Technology Decision Support System for the EBW Process* //Proceedings of the Computational Methods in Systems and Software. – Cham : Springer International Publishing, 456-466 (2022)
8. V. Kukartsev et al. *Methods and Tools for Developing an Organization Development Strategy* //2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). – IEEE, 1-8 (2022)
9. I. S. Masich, V. S. Tynchenko, V. A. Nelyub, V. V. Bukhtoyarov, S. O. Kurashkin, A. P. Gantimurov & A. S. Borodulin. Electronics, **11(24),** 4150 (2022)
10. I. A. Barantsov et al. Sensors, **23(2),** 582 (2023)
11. V. V. Bukhtoyarov et al. Electronics **12(1),** 215 (2023)
12. O. D. Repinskiy et al. Journal of Physics: Conference Series**. 1728(1)** 012032 (2021)
13. A. Rassokhin et al. Magazine of Civil Engineering, **109(1)** 10913 (2022)
14. A. Shutaleva et al. Education Sciences **12(5)** 324 (2022)
15. E. A. Efremenkov et al. Applied Sciences, **12(1)** 5 (2021)
16. A. Shutaleva et al. Sustainability, **14(1)** 250 (2021)
17. V. O. Gutarevich et al. Applied Sciences, **13 (8),** 4671 (2023)
18. B. V. Malozyomov et al. Energies, **16 (13),** 4907 (2023)
19. D. M. Strateichuk et al. Crystals, **13 (5),** 825 (2023)
20. B. V. Malozyomov et al. Energies, **16 (9),** 3909 (2023)
21. B. V. Malozyomov et al. Energies **16 (11)** 4276 (2023)
22. I. S. Masich, V. S., Tyncheko, V. A. Nelyub, V. V. Bukhtoyarov, S. O. Kurashkin & A. S. Borodulin. Computation, **10(10),** 185 (2022)
23. E. A. Kosenko and V. A. Nelyub. Matrix Polym. **15** 240(2) (2022)
24. V. A. Nelyub and I. A. Komarov. Metall, **202113** 1696(9) (2022)
25. V. A. Nelyub, S. Y. Fedorov, G. V. Malysheva and A. A. Berlin. Fibre Chem, **534** 53 252(7) (2022)
26. G. S. Sokolov, K. M. Shakirov and V. A. Nelyub. J. Phys. Conf. Ser. **1990** 012043 (2021)
27. V. Nelyub, S. Fedorov and Y. Klimovich. J. Phys. Conf. Ser. **1990** 012078 (2021)
28. K. M. Sharikov, G. S. Sokolov and V. A. Nelyub. J. Phys. Conf. Ser. **1990** 012044 (2021)
29. V. A. Nelyub. Polym. Sci. Ser. D 2021 **142** 14 260–4 (2021)
30. V. A. Nelyub, S. Y. Fedorov and G. V. Malysheva. Mater. Appl. Res. **124** 12 1037–41 (2021)
31. M. A. Orlov, V. A. Nelyub, A. N. Kalinnikov and A. S. Borodulin. J. Phys. Conf. Ser. **1990** 012042 (2021)
32. V. A. Nelyub. J. Phys. Conf. Ser. **1990** 012071 (2021)